

**Relatório de estágio submetido para obtenção do grau de Mestre
em Gestão Ambiental**



**Comparação da Modelação da Digestão
Anaeróbia da ETAR do Choupal através de
Redes Neurais Artificiais e ADM1**

Nádia Monteiro Amorim

Coimbra

Outubro 2016

Relatório de estágio submetido para obtenção do grau de **Mestre
em Gestão Ambiental**

Comparação da Modelação da Digestão Anaeróbia da ETAR do Choupal através de Redes Neurais Artificiais e ADM1

Orientador: Engenheiro Paulo Jorge Carvalho Leitão

Co-orientador: Engenheiro Roberto António Ferreira Barbosa

Co-orientador: Professor José Manuel Gonçalves

Nádia Monteiro Amorim

Coimbra

Outubro 2016

Agradecimentos

Ao meu professor e co-orientador José Manuel Gonçalves por todo o apoio e orientação prestados ao longo da realização deste trabalho, bem como, durante todo o mestrado que esteve sempre pronto para me ajudar em qualquer situação.

Ao meu orientador e co-orientador pela ajuda e orientação durante este estágio, especialmente ao Roberto Barbosa pelo apoio ao longo da elaboração deste trabalho estando sempre pronto para me esclarecer qualquer dúvida.

À Engenheira Joana Brandão do Polo Ria (Aveiro) e ao Engenheiro Hugo Sabino do CTGA, pelo apoio prestado para a compreensão dos softwares utilizados.

Aos meus colegas de gabinete, que sem dúvida tornaram a minha estadia na empresa muito mais animada, um especial agradecimento ao Paulo Oliveira pelas saídas para o terreno e por todo o apoio prestado, agradeço igualmente à Engenheira Isabel Quintaneiro por me ter convidado para participar na realização das colheitas onde aprendi muito. Agradeço também ao Engenheiro Pina Barreto pelo o cuidado e preocupação desde do primeiro dia de estágio. Ao Vasco e à Joana pelas muitas risadas no gabinete que tornavam o trabalho mais fácil.

Aos restantes trabalhadores da AdCL, que foram sempre muito simpáticos que me receberam de abraços abertos, permitindo assim uma rápida e fácil integração.

À minha amiga Sofia Serafim, que sem dúvida mesmo longe me ajudou na elaboração deste trabalho mas também ao longo do mestrado esteve sempre presente, sem ela teria sido difícil chegar até aqui.

À minha amiga e colega de estágio Joana Fortunato, foi realmente um apoio durante este tempo de estágio nunca me deixando desistir e por toda ajuda na realização deste relatório que sem dúvida foi muito importante.

Ao meu grupo de amigos pelo apoio e amizade nestes últimos anos, pelos momentos passados juntos, muita festa e risada que fazem parte desta vida académica.

À minha família, especialmente à minha irmã que estiveram sempre apoiar me ao longo destes anos.

Resumo

O aumento dos últimos anos de Estações de Tratamento de Águas Residuais (ETAR) tem sido benéfico para o meio ambiente, pois as ETAR's têm a função de receber e tratar as águas residuais de forma a serem devolvidas ao meio ambiente em condições ambientalmente seguras evitando assim a poluição.

Contudo as ETAR's acarretam muitos custos energéticos, posto isto, tem vindo a ser desenvolvidas soluções para colmatar estes custos.

Com este trabalho pretende-se otimizar o processo da Digestão Anaeróbia (DA) para se conseguir aumentar a produção de biogás que por sua vez é convertido em energia, contribuindo assim para a rentabilidade da exploração da ETAR. Nos últimos tempos as Redes Neurais Artificiais (RNA) têm sido das soluções mais usadas neste tipo de processo da DA, pois são modelos matemáticos computacionais, sendo capazes de prever valores quando novos casos lhes são apresentados, têm características de desempenho inspirados no cérebro humano. Os modelos neuronais artificiais têm como fonte de inspiração as redes neuronais biológicas.

Para o desenvolvimento das RNA foi utilizado o software NeuralTools da Palisade, esta consiste em quatro fases distintas, a preparação de dados, o treino, o teste e a previsão da rede.

O estudo de caso incide nos dois digestores da ETAR do Choupal da AdCl, empresa onde foi realizado o estágio curricular que originou este trabalho. Para o presente estudo foram utilizadas variáveis do tratamento de lamas da ETAR, como $\% MS_{LM}$, a $\% MO_{LM}$, que entra no digestor I, $Q_{entr.(m^3/dia)}$, o $Q_{biogás(m^3/dia)}$, a $\% MS$, a $\% MO$, a $T(^{\circ}C)$, o pH, a alcalinidade, e os ácidos gordos voláteis de cada um dos digestores. Foram treinadas e testadas várias redes até se obter a melhor rede para de seguida ser utilizada para prever os valores de biogás. Houve dificuldades na definição das quantidades e disponibilidade de alguns dados prejudicando assim as redes, não conseguindo prever valores semelhantes dos valores reais do processo da DA, mas ainda assim foi possível obter em alguns valores essa igualdade. Verificou-se que as variáveis com maior impacto na modelação do processo anaeróbio são a $MO(\%)_1$, a $MO(\%)_{LM}$, a $T(^{\circ}C)_1$, o $Q_{entr.(m^3/dia)}$. Conclui-se com este estudo, que usando o modelo RNA consegue-se obter resultados positivos no processo de DA permitindo assim otimizar a produção de biogás.

Abstract

The increase in recent years of wastewater treatment plants (WWTP) has been a benefit to the environment, this is due to the fact that WWTP's have the function of receiving and treating the waste water to be returned to the environment in environmentally safe conditions thus avoiding pollution.

However the WWTP's entails numerous energy costs. Due to that reason and to address these costs, solutions have been developed.

This assignment intends to optimize the process of anaerobic digestion (AD) to increase the production of biogas which in turn is converted into energy, thus contributing to the profitability of exploration of the WWTP. In recent times the artificial neural networks (Ann) have been the most commonly used solutions in this type of process because they are mathematical models, capable to predict values when new cases are presented, they have performance characteristics inspired by the human brain. Artificial neural models have as a source of inspiration the biological neural networks.

For the development of Ann the NeuralTools Palisade software was used, this consists of four distinct phases, the data preparation, training, testing and prediction of network. The case study focuses on the two digesters of WWTP of Choupal of AdCl, company where the curricular internship was held which led to this work. For this study we used variables of sludge treatment of WWTP to % MS_{LM}, % MO_{LM}, entering the digester I, Q_{entr} (m³/day), the Q_{biogás} (m³/day), the % MS, % MO, T (°C), pH, alkalinity and volatile fatty acids in each of the digesters. Several networks were trained and tested until the best network was found, to then be used to predict the values of biogas. There were difficulties in defining the quantity and availability of some data, hindering nets, failing to predict similar values of the real values of the process but it was still possible to obtain in some values that equality. It was found that the variables with the greatest impact in shaping the anaerobic process are the MO (%)₁, MO (%)_{LM}, T (°C)₁, Q_{entr} (m³/day). This study concludes that using the RNA template can achieve positive results in the process, as a result, optimizing biogás production.

Índice

1. Introdução.....	1
1.1. Âmbito do Relatório de Estágio.....	1
1.2. Enquadramento Geral do tema.....	1
1.3. Objetivos.....	4
2. Apresentação da Empresa e do Processo da ETAR do Choupal	5
2.1. Grupo Águas de Portugal.....	5
2.2. Grupo Águas do Centro Litoral	5
2.3. ETAR do Choupal	8
3. Processo de Digestão Anaeróbia.....	15
3.1. Fases Microbiologia e Bioquímica da Digestão Anaeróbia.....	15
3.2. Vantagens e desvantagens da Digestão Anaeróbia.....	17
3.3. Requisitos Nutricionais.....	17
3.4. Fatores que influenciam a Digestão Anaeróbia	19
3.5. Fatores prejudiciais ao processo da Digestão Anaeróbia.....	21
3.6. Modos de operação da Digestão Anaeróbia	22
3.6.1. Digestor anaeróbio de alta carga	23
3.6.2. Digestor anaeróbio de alta carga com sedimentação.....	23
3.6.3. Digestor anaeróbio de alta carga de duas fases	24
3.7. Produção de Biogás	25
4. Caracterização do Modelo GPS-X da Hydromantis	27
4.1. Esquema do tratamento anaerobio através do modelo GPS-X.....	28
5. Redes Neurais Artificiais.....	29
5.1. Evolução Histórica.....	30
5.2. Analogia com sistemas biológicos.....	31
5.3. Princípios de funcionamento e descrição das RNA.....	33
5.4. Tipos de Redes Neurais Artificiais	36
5.5. Fases de desenvolvimento de uma RNA	38
6. Apresentação do Caso de Estudo	41
6.1. Enquadramento do Estudo	41
6.2. Metodologia Proposta	41
6.2.1. Apresentação do NeuralTools	41

6.2.2.	Preparação de dados	42
6.2.3.	Treino e Teste da Rede Neuronal	45
6.2.3.1.	Configuração da rede	46
6.2.3.2.	Tempo de execução	47
6.2.3.3.	Processo de Treino	49
6.2.3.4.	Relatório de Treino.....	50
6.2.4.	Previsão	54
6.3.	Desenvolvimento do Estudo	55
6.3.1.	Introdução	55
6.3.2.	Preparação dos dados.....	55
6.3.3.	Treino e Teste	57
6.3.3.1.	Primeiro conjunto de redes.....	59
6.3.3.2.	Segundo conjunto de redes.....	59
6.3.3.3.	Terceiro conjunto de redes	60
6.3.4.	Validação	61
6.3.5.	Previsão	62
7.	Considerações finais.....	63
7.1.	Conclusões	63
7.2.	Sugestões para desenvolvimento futuros	65
8.	Referências Bibliográficas.....	66

Índice de Figuras

Figura 1-	Representação dos municípios abrangidos pela AdCL (Fonte: AdCL).....	6
Figura 2-	Representação dos municípios abrangidos pela AdCL (Fonte: AdCL).....	7
Figura 3-	Vista aérea da ETAR do Choupal (Adaptado da AdCL).....	8
Figura 4-	Esquema geral do processo da ETAR (Adaptado da AdCL).....	10
Figura 5-	Obra de entrada da ETAR do Choupal (Leonardo, 2012).	10
Figura 6-	Decantadores primários da ETAR do Choupal (Adaptado da AdCL).....	11
Figura 7-	Leitos percoladores da ETAR do Choupal (Adaptado da AdCL).	11
Figura 8-	Digestores anaeróbios da ETAR do Choupal.	13
Figura 9-	Desidratação de lamas da ETAR do Choupal.....	13
Figura 10-	Gasómetro da ETAR do Choupal.	14

Figura 11- Queimador de gás da ETAR do Choupal.	14
Figura 12- Fases envolvidas na DA (Deublein, 2008).	16
Figura 13- Sistema de um digestor de alta carga (Turovskiy & Mathai, 2006)	23
Figura 144- Sistema de duas fasescom sedimentação (Turovskiy & Mathai, 2006).	24
Figura 155- Sistema de duas fases de alta carga (Turovskiy & Mathai, 2006).	24
Figura 16- Esquema do tratamento anaeróbio da ETAR do Choupal, desenhado no simulado GPS-X.	28
Figura 17- Constituição de uma célula neuronal.	32
Figura 18- Modelo de um nó de uma RNA (Matos, 2008).	33
Figura 19- Esquema das ligações e elementos presentes num PMC de uma camada oculta (Matos, 2008).	36
Figura 20- Exemplo de uma base de dados para utilizar com o NeuralTools.	42
Figura 21- Caixa de diálogo do gestor do conjunto de dados.	43
Figura 22- Definição do tipo de variável.	44
Figura 23- Conjunto de dados definidos a ser utilizados.	45
Figura 24- Caixa de diálogo do treino.	46
Figura 25- Caixa de diálogo da Configuração da rede.	47
Figura 26- Caixa de diálogo do Tempo de execução.	48
Figura 27- Caixa de diálogo da Visão geral de treino.	49
Figura 28- Caixa de diálogo do Treinamento em Andamento.	50
Figura 29- Gráfico das diferenças residuais, Treinamento.	52
Figura 30- Gráfico relativo aos dados previstos comparados com os dados efetivos, Treinamento.	53
Figura 31- Gráfico referente às diferenças residuais, Teste.	53
Figura 32- Gráfico dos dados previstos comparando com os dados efetivos, Teste.	54
Figura 33- Gráfico do impacto das variáveis.	54
Figura 34- Gráfico da diferença residual, Treinamento.	80
Figura 35- Gráfico dos valores previstos comparados com os valores efetivos, Treinamento.	80
Figura 36- Gráfico com os valores residuais comparados com os valores efetivos, Treinamento.	81
Figura 37- Gráfico das diferenças residuais, Teste.	81
Figura 38- Gráficos com os valores residuais comparados com efetivos, Teste.	82
Figura 39- Gráfico do impacto das variáveis.	83

Figura 40- Gráfico de validação da melhor rede.	95
Figura 41- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 28°C.....	98
Figura 42- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 32°C.....	98
Figura 43- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 35°C.....	99
Figura 44- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 38°C.....	99
Figura 45- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 42°C.....	100

Índice de Tabelas

Tabela 1- Características principais da instalação da ETAR (Fonte: AdCL).....	9
Tabela 2- Capacidade calorífica do biogás e outros combustíveis (Abbasi, 2012).....	25
Tabela 3- Tabelas geradas no relatório de treino e teste automático do NeuralTools....	51
Tabela 4- Primeiras linhas da base de dados utilizada para treinar as redes neuronais..	71
Tabela 5- Tabela referente aos conjuntos de redes realizados no NeuralTools.....	75
Tabela 6- Tabela gerada do treinamento e teste automático da rede neuronal.....	79
Tabela 7- Tabela referente aos conjuntos de redes realizados no NeuralTools.....	85
Tabela 8- Tabela referente aos conjuntos de dados realizados no NeuralTools.....	89
Tabela 9- Tabela referente aos conjuntos de dados das melhores redes conseguidas pelo NeuralTools.	93
Tabela 10- Base de dados referentes aos valores utilizados para a validação.	95
Tabela 11- Base de dados referentes aos valores utilizados para a previsão da melhor rede.	97

Lista de Abreviaturas

AdCL- Águas do Centro Litoral

ADM1- Anaerobic Digestion Model No.1

ADP- Águas de Portugal

DA- Digestão Anaeróbia

EMQ- Erro Médio Quadrático

ETAR- Estação de Tratamento de Águas Residuais

GRNN- Generalized Regression Neural Network

MLFNN- MultiLayer Feed-forward Neural Network

PMC- Perceptrão Multicamadas

PNN- Probabilistic Neural Network

RNA- Redes Neurais Artificiais

TRH- Tempo de Retenção Hidráulica

1. Introdução

1.1. Âmbito do Relatório de Estágio

O seguinte relatório apresenta o trabalho realizado no âmbito da unidade curricular Estágio Profissionalizante, do Mestrado em Gestão Ambiental da Escola Superior Agrária de Coimbra realizado na ETAR do Choupal a qual pertence às Águas do Centro Litoral, do Grupo Águas de Portugal.

O estágio teve início no dia 7 de Março de 2016 e terminou no dia 7 de Setembro de 2016, e foi lançado com o tema: “Comparação da Modelação da digestão anaeróbia da ETAR do Choupal através de Redes Neurais Artificiais e ADM1”.

A proposta do tema tem em vista dar continuidade a uma abordagem que tem vindo a ser realizada em estágios realizados nos anos letivos anteriores, de aplicação de redes neurais artificiais na otimização do processo de digestão anaeróbia (DA) de Estações de Tratamento de Águas Residuais (ETAR) da AdCL, empresa do Grupo Águas de Portugal (AdP).

Para o desenvolvimento do presente trabalho, foram utilizados os softwares Microsoft Excel, da Microsoft, NeuralTools, da Palisade, e o GPS-X da Hydromantis que se encontram disponíveis na empresa AdCL.

Pretende-se, assim, encontrar o modelo que melhor simula o comportamento da digestão anaeróbia a fim de prever valores de produção de biogás.

1.2. Enquadramento Geral do tema

Uma das problemáticas mais preocupantes que se tem vindo a notar um pouco por todo o mundo é o aumento da poluição associado ao desenvolvimento socioeconómico da população, por isso, o Homem teve a necessidade de procurar soluções ambientais economicamente viáveis e práticas (Carreira, 2014).

Por essa razão, nos últimos anos tem havido um aumento significativo de Estações de Tratamento de Água Residuais (ETAR) que tem como principal função receber e tratar as águas residuais, de forma a serem devolvidas ao meio ambiente em condições ambientalmente seguras para prevenir a poluição ambiental e evitar as consequências que

os resíduos possam ter para a saúde pública, dado os microrganismos patogénicos e parasitas existentes (Águas do Centro Litoral, 2016).

Consequentemente com o aumento da população o uso dos combustíveis fósseis tem vindo a aumentar, pois o consumo é maior que a produção. Tanto a produção como a utilização de combustíveis fósseis levantam preocupações ambientais e problemas para a saúde pública. Para além, verificamos um aumento do preço dos combustíveis fósseis e da poluição ambiental, assim o Homem está a procurar soluções energéticas alternativas. Entre essas energias a que mais se tem destacado nos últimos anos é o aproveitamento da biomassa, visto que permite a recuperação do seu conteúdo energético na forma de biogás (Silveira, 2009).

O Biogás é uma mistura gasosa combustível, produzida através da digestão anaeróbica (DA), ou seja, pela biodegradação da matéria orgânica através da ação de bactérias na ausência de oxigénio (Freire, 2003).

As vantagens da produção de biogás sobre outro tipo de energias renováveis existentes no mercado, relaciona-se com o facto de conter um poder calorífico da mesma ordem de grandeza da do gás natural a um menor custo, sendo produzido através de matéria biodegradável considerada como resíduo, contribuindo para a redução de emissão de gases de efeito de estufa para a atmosfera. Por outro lado, a energia final do biogás pode ser aplicada em diferentes áreas de intervenção, tais como a combustão direta e produção de calor, geração de eletricidade e integração na rede de gás natural, utilização como combustíveis para veículos e, em última análise, a aplicação em pilhas de combustíveis (Silveira, 2009).

A digestão anaeróbia é um processo bioquímico que consiste na degradação de matéria orgânica na ausência de oxigénio molecular. Este é um dos sistemas mais utilizados na estabilização de lamas nas ETAR, uma vez que, do seu processo, parte da matéria orgânica volátil é convertida numa mistura de metano (CH_4) e dióxido carbono (CO_2) que pode ser aproveitada na produção de energia térmica e elétrica, evitando a sua libertação para a atmosfera (Metcalf&Eddy, 2003).

Dividindo-se em quatro principais fases: hidrólise, acidogénese (ou fermentação), a acetogénese e a metanogénese. Na primeira fase, ocorre a hidrólise das macromoléculas (compostos orgânicos complexos como as proteínas, hidratos de carbono e lípidos), na fase acidogénea ocorre a produção de compostos orgânicos solúveis e ácidos orgânicos de cadeia curta e na fase da acetogénese forma-se o acetado. Na última fase,

metanogénese, ocorre a conversão dos ácidos orgânicos voláteis em metano e dióxido de carbono formando então o biogás.

A escolha entre um processo de digestão aeróbia ou anaeróbia é efetuada tendo em conta os custos associados, a eficácia e a flexibilidade do processo. Dada a importância sobre a conservação de energia, a recuperação e uma reutilização benéfica dos biosólidos resultantes das águas residuais, o processo mais viável é o da digestão anaeróbia, visto que, este possibilita a produção de biogás que permite satisfazer algumas das necessidades energéticas da ETAR (Metcalf&Eddy, 2003).

Nos últimos anos, os estudos feitos em torno da DA têm contribuído para o conhecimento mais profundo das fases mais sensíveis do processo, por forma a conseguir uma melhor estabilidade e aproveitamento dos digestores anaeróbios.

Recentemente têm sido aplicadas às Redes Neurais Artificiais este tipo de processo para compreender melhor o seu funcionamento e assim otimizar a produção de biogás.

Uma RNA é um sistema de processamento de informação que possui algumas características de desempenho em comum com as redes neurais biológicas. Os modelos neurais artificiais têm como principal fonte de inspiração as redes neurais biológicas (Palisade, 2015).

Este tipo de modelação ainda é muito recente, ou seja, muitos não conhecem que esta ferramenta possa ser utilizada em várias áreas, como por exemplo na Engenharia e Ambiente, ao nível das previsões da precipitação diária, modelação de sistemas de abastecimento de água, ou no setor financeiro, análise e pesquisa de mercado, avaliação de créditos, entre outras (Carreira, 2014).

Outro tipo de modelação que se pode considerar é o GPS-X que é um programa que consiste na simulação dinâmica ou em regime permanente do tratamento de águas residuais domésticas e industriais usado no dimensionamento e exploração de ETAR com o objetivo de aumentar a eficiência de operação e reduzir respetivos custos, avaliar a máxima capacidade dos processos, avaliar estratégias de controlo e de beneficiação das ETAR face a diferentes cenários de exploração e treinar operadores (Ferreira F. M., 2012).

Utilizando estas duas ferramentas de simulação espera-se conseguir otimizar o processo da DA, aumentando assim a produção de biogás satisfazendo as necessidades energéticas da ETAR.

1.3. Objetivos

O principal objetivo na realização deste projeto consiste na comparação da modelação da digestão anaeróbia da ETAR do Choupal através de RNA e ADM1, conseguindo assim obter uma modelação matemática da produção de biogás através de outputs e inputs.

Para tal definiram-se os seguintes objetivos específicos:

- Relacionar algumas variáveis explicativas com a produção de biogás, definindo aquelas que têm maior impacto no processo de digestão anaeróbio;
- Aplicar RNA e ADM1 ao processo de DA de forma a desenvolver um modelo que permita otimizar a produção de biogás;
- Produzir um modelo de simulação da produção de biogás, com fim de apoiar a gestão do processo de digestão.

Desta forma, é possível verificar até que ponto os modelos de simulação de produção de biogás podem ser fiáveis para que seja possível otimizar o processo da digestão anaeróbia e por consequência aumentar a produção de biogás, ajudando na gestão de toda a ETAR.

2. Apresentação da Empresa e do Processo da ETAR do Choupal

2.1. Grupo Águas de Portugal

O Grupo Águas de Portugal (AdP) desenvolve atividades no domínio do abastecimento de água e do saneamento de águas residuais. Tem como missão conceber, construir, explorar e gerir Sistemas de Abastecimento de Água e de Saneamento de Águas Residuais, num quadro de sustentabilidade económica, financeira, técnica, social e ambiental, desenvolvendo um grupo empresarial português forte e de elevada competência, capaz de responder, com eficácia, aos grandes desafios que atualmente se colocam no setor do ambiente. Em parceria com os municípios, presta serviços a cerca de 80% da população portuguesa.

Nos serviços de abastecimento de água e saneamento de águas residuais em “alta”, que englobam as atividades de captação, tratamento e transporte até aos reservatórios municipais da água para consumo público, e do transporte, tratamento e devolução no meio recetor em condições ambientalmente seguras das águas residuais urbanas, as empresas do Grupo AdP abrangem 234 municípios.

O Grupo é ainda responsável pelos serviços de abastecimento e saneamento em “baixa”, diretamente ao consumidor, em alguns municípios, servindo cerca de 1,2 milhões de pessoas (Águas de Portugal, 2016).

2.2. Grupo Águas do Centro Litoral

O Sistema Multimunicipal de Abastecimento de Água e de Saneamento do Centro Litoral de Portugal foi criado a 29 de maio, pelo Decreto-Lei n.º 92/2015, e resulta da fusão do sistema multimunicipal de saneamento da ria de Aveiro, do sistema multimunicipal de saneamento do Lis, e do sistema multimunicipal de abastecimento de água e de saneamento do Baixo Mondego-Bairrada. Abrange a captação, o tratamento e o abastecimento de água para consumo público e a recolha, o tratamento e a rejeição de efluentes domésticos e urbanos, e a receção de efluentes provenientes de limpeza de fossas sépticas, que cumpram o disposto no regulamento de exploração e serviço relativo à atividade de saneamento de águas residuais em vigor no sistema, os respetivos tratamento e rejeição, que deve ser realizada de forma regular, contínua e eficiente.

A AdCL abrange cerca de 29 municípios nomeadamente Águeda, Albergaria-a-Velha, Ansião, Arganil, Aveiro, Batalha, Cantanhede, Coimbra, Condeixa-a-Nova, Espinho, Estarreja, Góis, Ílhavo, Leiria, Lousã, Marinha Grande, Mealhada, Mira, Miranda do Corvo, Murtosa, Oliveira do Bairro, Ourém, Ovar, Penacova, Penela, Porto de Mós, Santa Maria da Feira, Vagos e Vila Nova de Poiares (Águas do Centro Litoral, 2016).

Abastecimento de água

A AdCL tem como objetivo satisfazer as necessidades da população da região ao nível da quantidade e qualidade da água, abastecendo os municípios de Ansião, Arganil, Coimbra, Condeixa-a-Nova, Góis, Leiria, Lousã, Mealhada, Miranda do Corvo, Penacova, Penela, Vila Nova de Poiares, aos quais corresponde uma população de 339 mil habitantes;

A Empresa dispõe das seguintes infraestruturas: 11 captações, 4 estações de tratamento de água, 256 condutas (km), 31 estações elevatórias e 52 reservatórios (Águas do Centro Litoral, 2016).

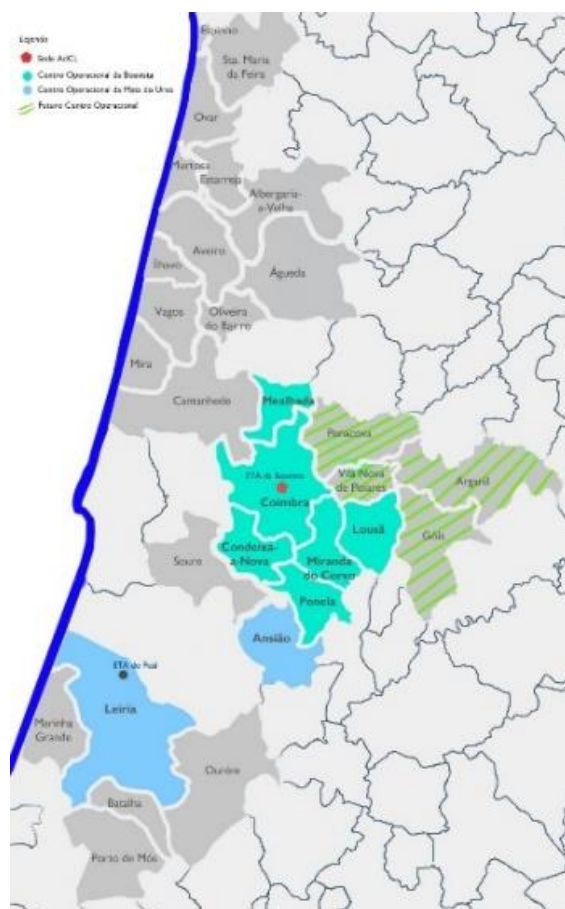


Figura 1- Representação dos municípios abrangidos pela AdCL (Fonte: AdCL).

Saneamento de Águas Residuais

A AdCL serve os municípios de Águeda, Albergaria-a-Velha, Aveiro, Batalha, Cantanhede, Estarreja, Espinho, Ílhavo, Marinha Grande, Mira, Murtosa, Oliveira do Bairro, Ovar, Ourém, Porto de Mós, Santa Maria da Feira, Soure, Vagos, representando 750 mil de habitantes-equivalentes.

Em relação às infraestruturas de saneamento, a empresa detém as seguintes: 67 estações de tratamento de águas residuais, 849 emissários (km), 158 estações elevatórias e 2 emissários submarinos (Águas do Centro Litoral, 2016).

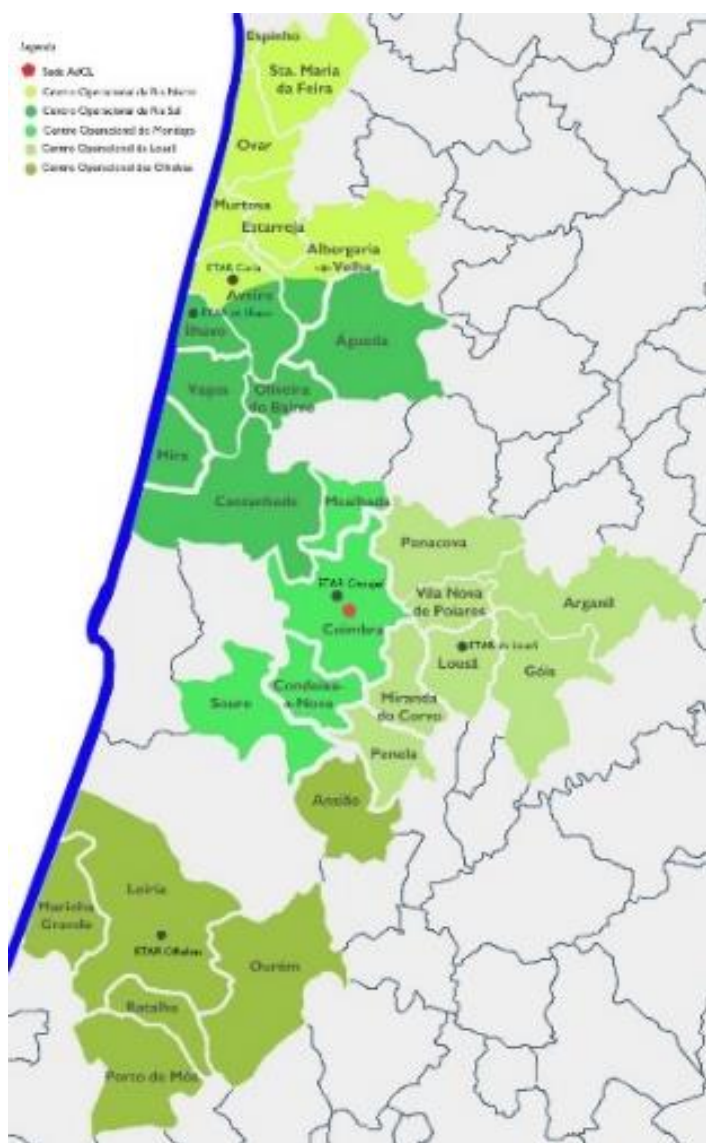


Figura 2- Representação dos municípios abrangidos pela AdCL (Fonte: AdCL).

2.3. ETAR do Choupal

A estação de tratamento de águas residuais do Choupal está em funcionamento desde 1992 e recebe os esgotos da zona norte e uma parte da zona sul do concelho de Coimbra. É a maior ETAR da região, tem uma capacidade de atendimento para aproximadamente 213 mil habitantes, e de tratamento de 36 000 m³/dia. Funciona em sistema de duas vias independentes, ou seja, se uma delas tiver algum problema de funcionamento, o tratamento não será interrompido por possuir uma segunda via de funcionamento.

Na seguinte figura mostra-se uma vista aérea onde se pode visualizar a disposição dos equipamentos na ETAR do Choupal.



Figura 3- Vista aérea da ETAR do Choupal (Adaptado da AdCL).

Legenda:

- 1-Obra de entrada
- 2- Decantadores Primários
- 3- Leitos percoladores
- 4- Decantadores secundários
- 5- Espessador
- 6- Digestores anaeróbios
- 7- Gasómetros
- 8- Oficina
- 9- Centrífuga
- 10- Leitos de secagem

De seguida apresenta-se uma breve descrição das principais características da instalação da ETAR do Choupal.

Tabela 1- Características principais da instalação da ETAR (Fonte: AdCL).

Caraterísticas principais da instalação da ETAR do Choupal	
Pré Tratamento	3 Grelhas de limpeza mecânica 2 Desengorduradores/desarenadores com A=24m
Tratamento Primário	2 Decantadores primários, d=37m
Estação Elevatória	2 Grupos Q=250L/s, P=45kW 2 Grupos Q=500L/s, P=90kW
Tratamento Biológico	4 Leitos percoladores, d=36m 2 Decantadores secundários, d=37m
Tratamento de Lamas	1 Espessador gravítico, d=18m 2 Digestores anaeróbios, d=20m, V=3 000m ³
Gasómetro	2 Gasómetros, d=14,5m, V=1 000m ³
Aquecimento de Lamas	2 Caldeiras com potência calorífica de 640 000 kcal/h
Desidratação de Lamas	Centrifuga
Potência Instalada	1 000 kVA
Destino final do Efluente	Rio Mondego

No processo de tratamento desta ETAR podem destacam-se as três linhas de tratamento, a linha de água, a de lamas e a de gás, na figura seguinte está representado o esquema geral e posteriormente a explicação de cada etapa detalhada.

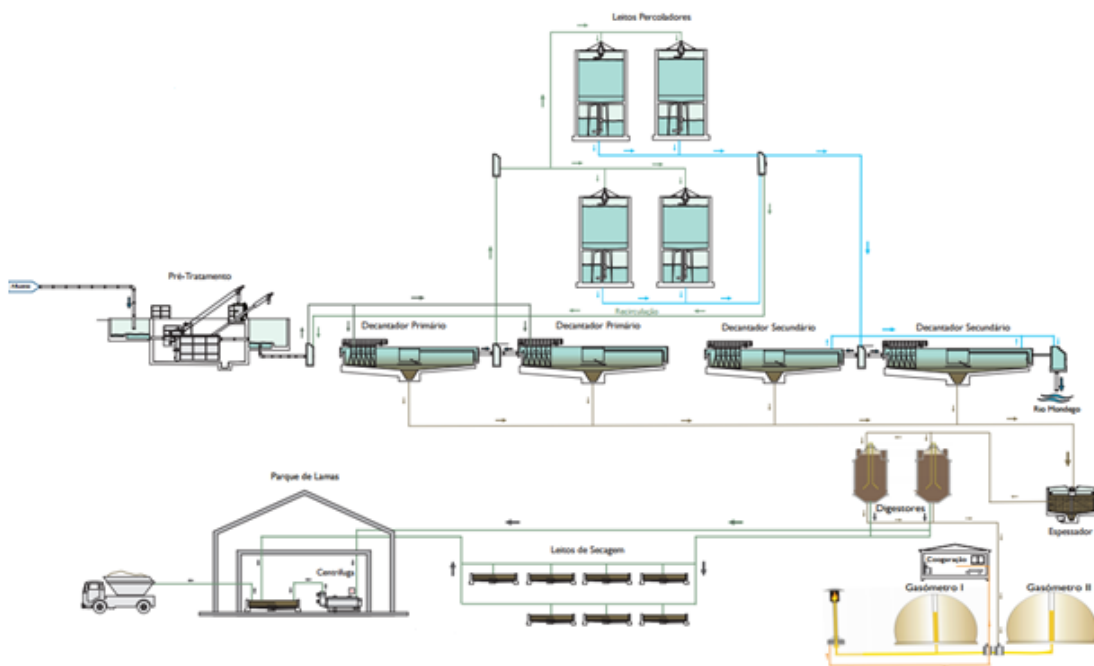


Figura 4- Esquema geral do processo da ETAR (Adaptado da AdCL).

Linha de Água

- **Pré-Tratamento – Obra de entrada**

O Pré-Tratamento é a entrada do efluente, onde ocorre o processo de gradagem, em que são removidos os sólidos de maiores dimensões com auxílio de grelhas mecânicas. Os sólidos são armazenados em contentores e reencaminhados para o aterro. Ainda no processo, são removidas ainda as areias e gorduras, estas são separadas pela introdução de ar e as areias são extraídas, após sedimentação, por bombagem (Leonardo, 2012).



Figura 5- Obra de entrada da ETAR do Choupal (Leonardo, 2012).

- **Decantação Primária**

A Decantação Primária é o efluente vindo da obra de entrada, onde entra num compartimento de divisão de caudais seguindo para a decantação primária, esta é constituída por dois decantadores a funcionar em paralelo, de planta circular, providos de ponte raspadora e do sistema de remoção de sobrenadantes. A matéria sedimentada é chamada de lama primária e acumula-se no fundo cónico à medida que a ponte a arrasta, sendo de seguida encaminhada graviticamente para a fossa de lamas primária (Leonardo, 2012).



Figura 6- Decantadores primários da ETAR do Choupal (Adaptado da AdCL).

- **Tratamento Biológico**

Dos decantadores primários, o efluente passa para os leitos percoladores nos quais ocorre o tratamento biológico. Os quatro leitos percoladores são dotados de distribuidores rotativos e de canais de ventilação. O efluente a tratar é elevado a partir de uma fossa com quatro grupos de bombagem. À medida que o efluente escorre pelas pedras forma-se o filme biológico que serve de suporte aos microrganismos e que em contacto com o ar removem a matéria orgânica dissolvida nas águas residuais (Leonardo, 2012).



Figura 7- Leitos percoladores da ETAR do Choupal (Adaptado da AdCL).

- **Decantação Secundária**

A Decantação Secundária tem como principal função a clarificação do efluente e é constituída por dois decantadores de planta circular, a funcionar em paralelo, providos de ponte raspadora e sistema de remoção de sobrenadantes. A lama secundária acumula-se no fundo cónico à medida que a ponte a arrasta e de seguida encaminhada graviticamente para a fossa de lamas secundárias. O efluente depurado, é recolhido por transbordo sobre o descarregador periférico e é conduzido até ao rio Mondego através de uma conduta (Leonardo, 2012).

Linha de Lamas

A formação das lamas mistas resultam da mistura no fosso de lamas resultantes do tratamento da fase líquida.

Antes de seguirem para os digestores anaeróbios, as lamas passam por um espessador, no entanto, no Choupal este equipamento encontra-se desativado, ou seja, as lamas são bombeadas diretamente da fossa de lamas para o tanque de alimentação do digestor.

- **Digestão de Lamas**

Ainda de acordo com (Leonardo, 2012) os digestores encontram-se a operar em simultâneo, em regime em série, a diferença na operação dos digestores encontra-se principalmente na agitação e alimentações realizadas. O digestor primário é agitado quase continuamente, em curtos períodos ao longo de todo dia, sendo alimentando cerca de 4 a 5 vezes ao dia. O digestor secundário é agitado apenas uma vez ao dia, e recebe apenas uma alimentação que corresponde aproximadamente à mesma quantidade das 4 a 5 alimentações de digestor primário. São introduzidas diariamente, em modo descontínuo lamas mistas, pelo que são previamente aquecidas no permutador, antes de entrarem no primeiro digestor. O mesmo acontece com o segundo digestor, em que as lamas do primeiro digestor, são previamente aquecidas num segundo permutador. No digestor secundário são formadas as lamas de clarificado e as lamas da saída, ambas retiradas diariamente de modo a manter o nível constante.

Na prática, a operação em estado estacionário nem sempre é conseguida, pois o sistema é constantemente alvo de perturbações que afetam o nível e a diluição das lamas e por consequência a estabilidade da produção de biogás. Globalmente, os digestores apresentam um volume mínimo e máximo de enchimento que deve ser respeitado. A

altura mínima assegura que o biogás não saia pela saída das lamas, e o limite máximo garante que as lamas não saiam pela saída do biogás.



Figura 8- Digestores anaeróbios da ETAR do Choupal.

- **Desidratação de Lamas**

As lamas provenientes dos digestores anaeróbios são desidratadas recorrendo a uma centrífuga sendo posteriormente enviadas para valorização em terrenos agrícolas.



Figura 9- Desidratação de lamas da ETAR do Choupal.

Linha de Gás

O biogás produzido na ETAR pode ser utilizado na agitação dos digestores, para alimentar as caldeiras e as turbinas (biogás consumido), armazenado (gasómetros), ou queimado sem aproveitamento energético. O biogás usado na agitação pode ser captado diretamente dos digestores, sendo injetado na zona inferior dos digestores, através de seis

difusores, criando correntes convectivas de baixo para cima. O aproveitamento energético tem elevado interesse económico, mas neste caso o biogás tem de ser previamente submetido a um tratamento de forma a tornar viável a sua utilização. O restante biogás não utilizado é armazenado nos gasómetros mas antes de ser armazenado e queimado, o biogás é conduzido a um filtro de condensados onde lhe é removido o vapor de água. Para ser utilizado na produção de energia elétrica (turbinas) e aquecer a água das caldeiras, é depurado num filtro de finos, onde são removidas poeiras e condensados (Leonardo, 2012).



Figura 10- Gasómetro da ETAR do Choupal.

Quando não existe a capacidade de armazenar mais biogás nos gasómetros e em situações de emergência, utiliza-se um queimador de gás.



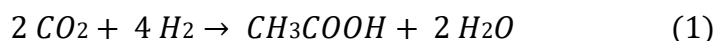
Figura 11- Queimador de gás da ETAR do Choupal.

3. Processo de Digestão Anaeróbia

3.1. Fases Microbiologia e Bioquímica da Digestão Anaeróbia

A digestão anaeróbia de compostos orgânicos pode ser dividida em quatro fases: hidrólise, acidogénese, acetogénese e metanogénese.

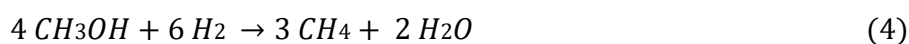
- **Hidrólise:** compostos orgânicos complexos, como lípidos, proteínas, ácidos nucleicos e polissacarídeos são hidrolisados por ação de enzimas extracelulares (hidrolases) que os convertem em compostos mais simples, tais como, aminoácidos, monossacarídeos e ácidos gordos de cadeia longa capazes de atravessar a membrana celular dos microrganismos para serem convertidos. No entanto, o tempo de decomposição destas macromoléculas é variável: no caso dos polissacarídeos, a hidrólise pode ocorrer em poucas horas, ao contrário no caso das proteínas e lípidos, são necessários alguns dias. A lignocelulose e a lignina são compostos muito difíceis de decompor pelas bactérias, sendo a sua hidrólise muito lenta, pelo que o tempo de permanência não é suficiente para a sua decomposição. Em geral, a hidrólise pode ser considerada como uma das etapas limitantes da DA (Deublein, 2008).
- **Acidogénese:** etapa mediada pelas bactérias acidogénicas, responsáveis por utilizar compostos orgânicos produzidos na hidrólise para produzir cadeias pequenas de ácidos orgânicos, tais como, ácido acético, ácido propiónico, ácido butírico, ácido valérico, bem como, álcoois, hidrogénio e dióxido de carbono (Deublein, 2008).
- **Acetogénese:** os substratos produzidos na acidogénese são utilizados para produzir ácido acético, dióxido de carbono, hidrogénio
Está presente na Eq. (1) a conversão de dióxido de carbono e hidrogénio a ácido acético e água.



Estas bactérias produzem ainda H_2 , composto importante à sobrevivência das bactérias metanogénicas. No entanto, para proporcionar um ambiente adequado à sobrevivência das bactérias acetogénicas é necessário que a pressão parcial de hidrogénio seja reduzida,

o que é possível com o consumo deste substrato pelas metanogénias. Nesta fase, observa-se ainda a redução dos sulfatos a H^2 (Deublein, 2008).

- **Metanogénese:** considerada a etapa mais limitante do processo, onde bactérias específicas convertem principalmente ácido acético, hidrogénio e dióxido de carbono em metano. A conversão destes compostos é realizada por grupos diferentes de metanogénicas, as hidrogenotróficas (Eq.2), acetoclásticas (Eq.3), e metilotróficas (Eq.4).



A via mais comum de produção de metano é a partir do ácido acético (Eq.2), sendo que a partir de metanol e hidrogénio é a menos observada (Eq.3) (Deublein, 2008).

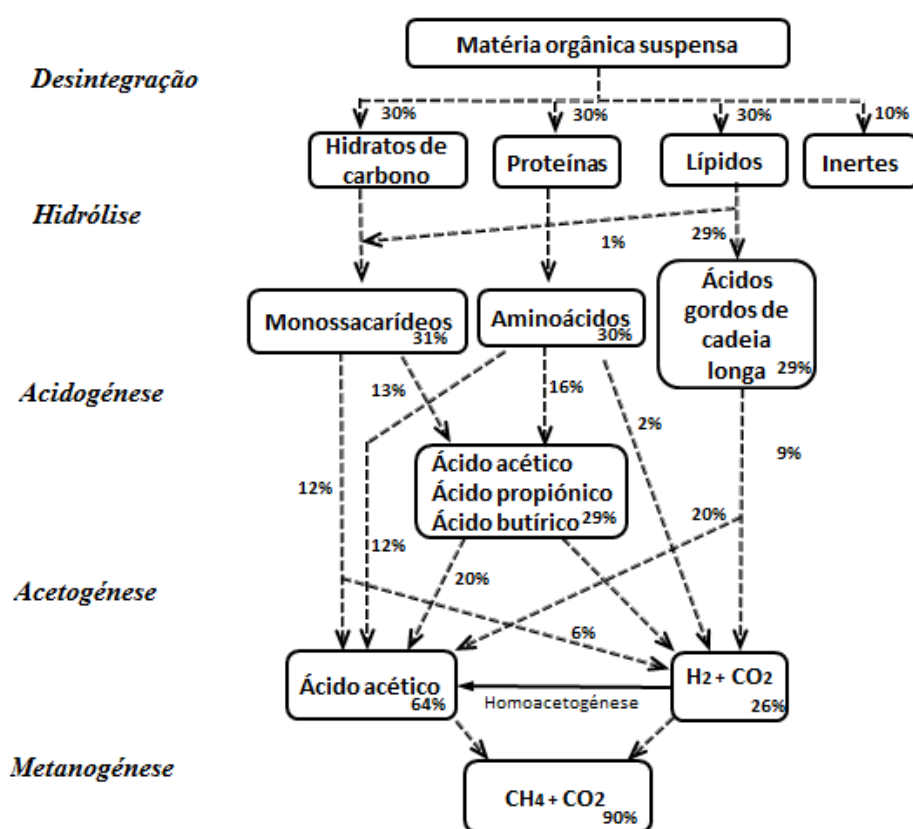


Figura 12- Fases envolvidas na DA (Deublein, 2008).

3.2. Vantagens e desvantagens da Digestão Anaeróbia

Segundo (Mendes, 2014) citado por (Fernandes, 2015) o processo da digestão anaeróbia tem muitas vantagens sendo por isso um dos processos mais utilizados para a estabilização das lamas, tais como:

- A produção de biogás, que pode ser utilizado para produzir energia elétrica para sustentar total ou parcialmente os equipamentos da ETAR ou mesmo para vender essa energia à rede elétrica. Para melhorar o funcionamento, a produtividade e até mesmo o desempenho pode fazer-se o aquecimento das lamas na digestão utilizando a energia térmica produzida na combustão do biogás;
- Redução da massa e volume total de lamas através da conversão em metano, dióxido de carbono e vapor de água. A produção de lamas digeridas pode ser utilizada como fertilizante e contribuir para o aumento da textura do solo por conterem nutrientes, como azoto, fósforo, matéria orgânica, entre outros.

Como em todos os processos, referindo (Turovskiy & Mathai, 2006) este também tem algumas desvantagens, como:

- São elevados custos de instalação e manutenção, devido às grandes dimensões dos digestores e equipamentos associados a este processo;
- Necessidade de monitorização e controlo do processo para prevenir anomalias, pois os microrganismos presentes neste processo são muito sensíveis;
- Digestores anaeróbios que não são agitados, ou cujo conteúdo não é homogeneizado estão na origem de sobrenadantes com alta concentração de sólidos suspensos, azoto e fósforo.

3.3. Requisitos Nutricionais

É necessário garantir a satisfação dos requisitos nutricionais para assegurar a eficiência e estabilidade do processo de degradação anaeróbia.

Os processos anaeróbios são conhecidos pela baixa produção celular, logo ocorre uma redução das necessidades em nutrientes. As concentrações necessárias encontram-se

diretamente relacionadas com a concentração de substrato a utilizar pelos microrganismos (Brandão, 2014).

- **Macronutrientes**

Os nutrientes como o azoto, o fósforo e o enxofre são precisos em maior quantidade durante o processo anaeróbio (Mara & Horan, 2003).

A razão carbono/azoto é importante no processo de digestão anaeróbia, uma vez que um destes componentes pode tornar-se limitante. O azoto é o que leva a que as bactérias tenham um crescimento adequado para o normal funcionamento de um digestor, já o carbono funciona como fonte de energia para os microrganismos existentes (Silveira, 2009; Carreira, 2014).

A fonte de azoto mais pretendida pelos microrganismos metanogénicos é o ião amónio, pois a sua presença faz aumentar o crescimento das bactérias referidas (Alves, 1998).

Em relação à fonte de fósforo, quer o ácido fosfórico quer o sal de fosfato são ótimos nutrientes para adicionar ao digestor. Os precipitados inorgânicos de fosfato podem funcionar, nos digestores anaeróbios, como fonte interna de fósforo.

O enxofre é outro macronutriente importante nas reações metanogénicas e são utilizados sulfuretos que podem levar à precipitação de alguns metais necessários no meio, como o ferro, níquel e cobalto (Speece, 1983; Carreira, 2014).

- **Micronutrientes**

O ferro, o cobalto e o níquel são compostos importantes para o bom funcionamento deste processo, mas em quantidades mais pequenas.

O ferro funciona como fonte de eletrões que intervém no metabolismo e está presente em citocromos e ferredoxinas (Alves, 1998; Carreira, 2014).

O cobalto afeta a metanogénese a partir do metanol pois estimula as bactérias intervenientes nesta fase que utilizam diretamente o metanol (Carreira, 2014).

Por último mas não menos importante, tem-se o níquel, que é um dos constituintes envolvidos nos passos terminais das reações metanogénicas. Nas membranas celulares existem elevados níveis de níquel, razão pela qual se encontra relacionado com a estabilidade estrutural (Brandão, 2014).

3.4. Fatores que influenciam a Digestão Anaeróbia

A digestão anaeróbia é um processo biológico complexo e sensível a alterações, tornando-o por vezes difícil de controlar. Desta forma, existem fatores que afetam a sua estabilidade:

- **Temperatura** é um parâmetro a ter em conta na digestão anaeróbia, quer do ponto de vista cinético quer do ponto de vista termodinâmico. Citado por (Carreira, 2014) a metanogénese é fortemente influenciada por este parâmetro, as taxas de degradação e dos coeficientes de crescimento são dependentes da temperatura. Na digestão anaeróbia existe uma variedade de gamas para o seu funcionamento:

- Gama psicrófila ($< 20\text{ }^{\circ}\text{C}$)
- Gama mesófila ($20\text{-}45\text{ }^{\circ}\text{C}$)
- Gama termófila ($> 45\text{ }^{\circ}\text{C}$)

As gamas termófila e mesófila, são aquelas em que há um crescimento máximo de microrganismos da fase metanogénica. A temperatura num digestor deve ser constante o máximo possível para evitar efeitos negativos para o crescimento dos microrganismos presentes (Alves, 1998 (Cavaleiro, 1999)).

O regime termófilo possui várias vantagens, tais como:

- Aumenta a solubilidade dos compostos orgânicos;
- Aumenta as taxas de reações químicas e biológicas dentro do digestor;
- Taxa de destruição das bactérias patogénicas também pode aumentar.

Mas também se encontram desvantagens ao utilizar temperaturas elevadas pois, uma vez que o processo se pode tornar instável, pode levar a um aumento da fração de amónia livre (NH_3) que é inibitória para o processo, bem como de AGV que necessita também de um maior consumo de energia para controlar a temperatura. Por outro lado, o processo mesófilo pode se tornar mais vantajoso pois não exige um controlo tão grande, sendo mais estável (Van Lier, 1995 (Brandão, 2014)).

- A **Alcalinidade** refere-se à resistência do meio a variações de pH, trata-se da medida representativa da capacidade tampão do meio. Na digestão anaeróbia os AGV é um produto intermediário que provocam uma elevada descida de pH do

meio, caso a alcalinidade não seja suficiente. Quando o poder tampão do efluente a tratar é fraco, é necessário corrigir o pH para isso pode-se adicionar um reagente alcalino (Cavaleiro, 1999; Brandão, 2014).

- O **pH** é muito importante na atividade enzimática de cada microrganismo, cada um deles possui um pH ótimo que se traduz numa máxima atividade. Sendo as bactérias metanogénicas, as responsáveis pela cinética global do processo, considera-se que a gama de pH ótimo para estes microrganismos encontra-se entre 6,6 a 7,6 (Cavaleiro, 1999; Silveira, 2009). O pH afeta a taxa de crescimento dos microrganismos bem como a sua morfologia e a sua estrutura, o que pode dificultar os fenómenos de adesão e floculação no digestor (Forage, 1985). Além disto, alguns compostos podem tornar-se tóxicos devido a grandes mudanças de pH, ou seja, para valores de pH baixos, os AGV livres podem causar inibição por ácidos fracos, enquanto para valores de pH elevados, a amónia livre pode causar inibição por bases fracas. Em ambos os casos geram-se formas não ionizadas que penetram a célula e aumenta a concentração intracelular de hidrogénio no meio, dadas as dissociações que provocam (Fukuzaki, 1990).
- **TRH** indica o tempo médio que um determinado volume de lamas permanece no digestor anaeróbio, é um dos parâmetros mais importantes para o dimensionamento de um sistema de tratamento anaeróbio pois determina o tempo de contacto entre o substrato e as bactérias.
Como as bactérias metanogénicas são de crescimento lento e são microrganismos determinantes na produção de biogás, o TRH deve ser sempre superior ao tempo de duplicação dessas bactérias, de modo a garantir uma degradação total da matéria (Tchobanoglous G. B., 1991), citado por (Fernandes, 2015).
- A **Agitação** tem como objetivo melhorar o contato entre o substrato do afluente com a biomassa no interior do digestor, evitando a variação de temperatura dentro do digestor e mantendo a mesma concentração de substrato. Para além disso, a agitação evita a formação de camadas de espuma superficiais e a deposição de lama no fundo do digestor (Rocha, 2013).

3.5. Fatores prejudiciais ao processo da Digestão Anaeróbia

Existem diversos tipos de microrganismos que prejudicam o processo da DA, estes podem levar à diminuição da velocidade de processo- **toxicidade** ou mesmo paragem do mesmo- **inibição** (Carreira, 2014).

Os compostos inibitórios podem ser classificados como (Cavaleiro, 1999):

- **Inibidores metabólicos:** causam inibição reversível, após o seu desaparecimento do meio a atividade do digestor volta ao normal;
- **Inibidores fisiológicos:** afetam os produtos formados ao longo da digestão, de forma que os seus efeitos continuam após o seu desaparecimento do meio;
- **Agentes bactericidas:** conduzem à morte da flora bacteriana.

Nestes compostos pode-se encontrar: amónia, sulfuretos, metais pesados, metais alcalinos e metais alcalino-terrosos. De seguida apresenta-se uma breve descrição destes:

- **Amónia** é um composto libertado durante a hidrólise através da degradação biológica da matéria azotada que se encontra presente essencialmente na forma de proteínas. Dependendo dos níveis de pH pode formar-se o ião amónio (NH_4^+) ou amónia livre (NH_3). Este último é mais tóxico, pois é a forma não ionizada e formada a pH mais elevados, tem a capacidade de penetrar nas paredes das células levando a uma desregulação do balanço de protões bem como à deficiência de potássio (Cavaleiro, 1999; Silveira, 2009). Já a sua forma ionizada tem um efeito contrário no processo pois a sua presença leva à produção de hidróxido que ao reagir com o dióxido de carbono forma bicarbonato, aumentando a capacidade tampão do processo (Silveira, 2009).

Quer o pH, a temperatura ou a concentração total de amónia, são fatores que interferem na concentração da amónia livre existente no digestor. Normalmente um aumento do pH ou da temperatura leva a um aumento da toxicidade, inibindo a atividade metanogénica, dado que aumenta a concentração da forma não ionizada comparando com a forma ionizada (Borja, Sanchez, & Weiland, 1996; Brandão, 2014).

- **Sulfuretos** são as formas inorgânicas de enxofre presentes na alimentação do digestor, principalmente o sulfato, são rapidamente convertidas pelas bactérias redutoras de sulfato para as formas reduzidas de sulfureto e sulfureto de hidrogénio, que são classificados como inibidores importantes da digestão anaeróbia (Brandão, 2014).

A redução do sulfato nos digestores anaeróbios numa maneira geral é um processo indesejável, visto que a formação de sulfureto causa muitos problemas como toxicidade, corrosão, odor, além de diminuir a qualidade e quantidade do biogás produzido. Segundo (Silveira, 2009) e citado por (Brandão, 2014) a produção de sulfuretos pode apresentar vantagens ao processo de digestão, como provocar a precipitação de metais pesados tóxicos (crómio, cobre e zinco), visto que os precipitados são bons precursores para a formação de grânulos, e aumentarem a alcalinidade no digestor derivado da produção de bicarbonato.

- **Metais Pesados** para a síntese celular dos microrganismos que estão envolvidos na DA são necessários os metais pesados. Existem limites de tolerância para o teor destes microrganismos que ao serem ultrapassados podem resultar num aumento da toxicidade e prejudicar o processo da digestão anaeróbia (Mendes, 2014) citado por (Fernandes, 2015).

O crómio, ferro, cobalto, cobre, zinco, entre outros podem ser considerados como metais pesados principais presentes na drenagem das águas residuais, e em determinadas concentrações podem se tornar um grande problema.

Contudo quanto maior o teor em sólidos menor a suscetibilidade do sistema entrar em inibição e a presença de sulfuretos provoca a precipitação destes compostos (Silveira, 2009).

- **Metais alcalinos e alcalino-terrosos** utilizando baixas concentrações de sódio, potássio, cálcio e o magnésio consegue-se o controlo de PH mas se são feitas adições excessivas podem ter um efeito inibitório no processo da DA.

O magnésio ou a combinação de cálcio e potássio podem reduzir a toxicidade provocada pelo sódio e consequentemente o efeito combinado de cálcio e sódio reduz a toxicidade provocada pelo potássio (Silveira, 2009; Cavaleiro, 1999).

3.6. Modos de operação da Digestão Anaeróbia

Os digestores de uma maneira geral podem trabalhar em modo contínuo e descontínuo. No caso de contínuo, tenta-se operar em regime estacionário, ou seja, os caudais de

entrada e saída devem ser iguais, de modo a manter o volume de lamas constante nos digestores. Em regime descontínuo, a alimentação e remoção da suspensão ocorrem em determinados tempos.

3.6.1. Digestor anaeróbio de alta carga

A digestão ocorre num único digestor, que é dotado de aquecimento por permutadores externos e agitação constante efetuada através da recirculação de biogás no interior dos digestores, como mostra na figura 13. O aquecimento por permutadores de calor externos é o método mais comum por causa da sua flexibilidade e facilidade de manutenção. Em relação à mistura encontrada no digestor, esta é importante para reduzir a estratificação térmica, melhorar o contacto das lamas brutas com a biomassa, reduzir a formação de espuma, entre outros (Turovskiy & Mathai, 2006) citado por (Leonardo, 2012). As lamas devem ser alimentadas continuamente ou em intervalos regulares, para manter as condições de estado estacionário no digestor e reduzir os picos de carga (Appels, Baeyens, Degreve, & Dewil, 2008). Estes tipos de digestores são tipicamente dimensionados para um tempo de retenção de sólidos de 10 a 20 dias, este processo é o mais utilizado e de fácil operação, podendo ser classificado mediante temperaturas na faixa mesófila ou termófila (Turovskiy & Mathai, 2006).

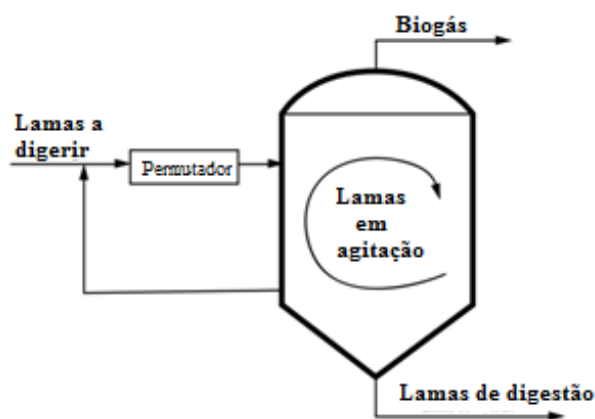


Figura 13- Sistema de um digestor de alta carga (Turovskiy & Mathai, 2006)

3.6.2. Digestor anaeróbio de alta carga com sedimentação

Este sistema é constituído por dois digestores, onde no primeiro digestor (digestor primário) é de alta carga e é dotado de agitação e aquecimento, onde ocorre a formação de biogás. No segundo digestor (digestor secundário) ocorre a separação sólido-líquido,

pelo que onde não existe aquecimento nem agitação. Este sistema é muitas vezes utilizado como um sistema que permite armazenar as lamas, permitindo a sua estratificação, auxiliando assim, o processo a jusante de desidratação, reduzindo os gastos com polímero utilizado e as perdas de biogás (Tchobanoglous G. B., 2003). Este último sistema não é o mais recomendado, pois tem mais custos comparativamente com apenas um digestor e não contribui no aumento da produtividade em biogás (Lettinga, 1995).

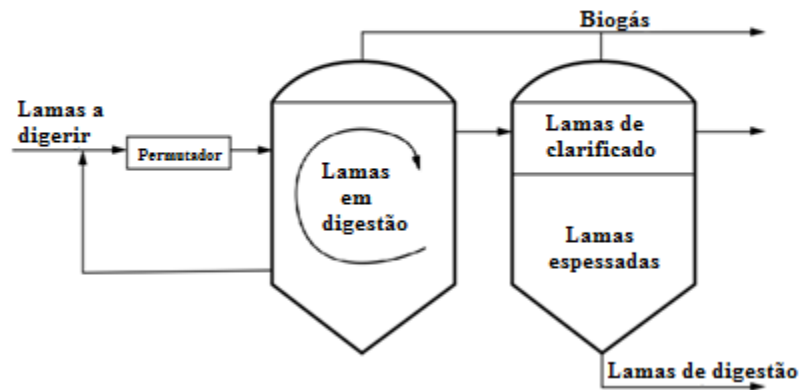


Figura 144- Sistema de duas fases com sedimentação (Turovskiy & Mathai, 2006).

3.6.3. Digestor anaeróbio de alta carga de duas fases

O sistema de duas fases de alta carga é utilizado para otimizar a produtividade de metano, conferindo condições ótimas em cada digestor para a atividade bacteriana, estando ambos sob agitação e aquecimento.

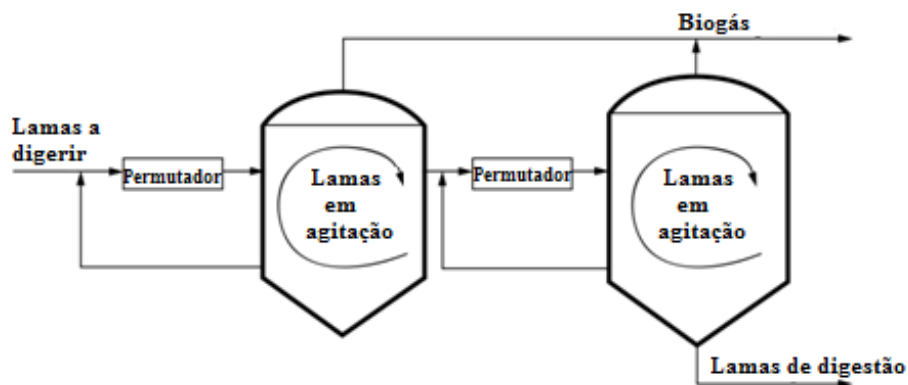


Figura 155- Sistema de duas fases de alta carga (Turovskiy & Mathai, 2006).

No sistema de duas fases ocorre a separação das fases da DA: fase ácida (digestor primário) e a fase da produção de biogás (digestor secundário). A vantagem deste processo é fornecer condições ótimas aos diferentes grupos de bactérias com vista a aumentar a produção de biogás. Assim, para o digestor primário que é projetado para um

tempo de retenção de 1 a 2 dias, para gamas mesófilas ou termófilas e valores de pH entre 5,2 a 6,3, ideal para a hidrólise e acidogénese. Por sua vez, o segundo digestor funciona para temperaturas mesófilas e para tempos de retenção de cerca de 10 dias, ocorre a acetogénese e metanogénese, onde se dá a formação de biogás.

Assim, pretende facilitar a degradação dos compostos (hidrólise), eliminar os microrganismos patogénicos e obter sólidos de maior qualidade.

3.7. Produção de Biogás

O biogás é um biocombustível gasoso constituído maioritariamente por metano (CH_4) dióxido de carbono (CO_2), contendo também, vestígios de compostos, tais como ácido sulfídrico (H_2S), oxigénio (O_2), amoníaco (NH_3). O biogás pode ser produzido em ETAR's, aterros sanitários e unidades de co-digestão, é constituído por 40 a 70% de CH_4 , 30 a 60% de CO_2 e contém vestígios de amoníaco, ácido sulfídrico e hidrogénio (Abbasi, 2012). O metano é o principal constituinte do biogás, o qual é um gás incolor, inodoro, com propriedades inflamáveis e contribui cerca de 21 vezes mais para o efeito de estufa comparativamente com o CO_2 .

O biogás tem um valor calorífico apreciável, e portanto, com aproveitamento energético. Para um teor de 55-65% de CH_4 , permite obter 5,5 kWh/m³ de biogás e uma massa volumica de 1,2 kg/m³ (Deublein, 2008). De seguida são apresentadas as capacidades caloríficas de diversos combustíveis comparativamente com a de biogás.

Tabela 2- Capacidade calorífica do biogás e outros combustíveis (Abbasi, 2012)

Combustível	Capacidade calorífica
Gás natural	8600 (kcal/m ³)
GPL	10800 (kcal/Kg)
Querosene	10300 (kcal/Kg)
Diesel	10700 (kcal/Kg)
Biogás	5000 (kcal/m ³)

O aproveitamento energético do biogás envolve geralmente a produção de energia elétrica, térmica e produção de biometano (gás natural), apesar de ter de ser previamente submetido a processos de limpeza e purificação.

Destino final das lamas de digestão e sua Legislação

No fim do tratamento das lamas, estas tem um destino final, mas é necessário cumprir determinadas exigências relativamente à qualidade das mesmas.

O principal destino das lamas é a **valorização na agricultura**, pelo que o seu aproveitamento consiste numa das aplicações com maior potencial, na medida em que podem ser utilizadas como fertilizante, devido ao seu pH mais alcalino e à presença de nutrientes que podem colmatar as necessidades nutricionais dos solos. No entanto a aplicação agrícola de lamas está condicionada por diversos fatores, nomeadamente as características dos solos e das lamas, etc.

Para que as lamas possam ser utilizadas para valorização agrícola, é fundamental que o tratamento efetuado permita cumprir todos os requisitos legais associados a esta prática. O Decreto-Lei nº 276/2009 estabelece o regime de utilização de lamas de depuração em solos agrícolas, promovendo a sua correta utilização, e transpõe para a ordem jurídica interna a Diretiva n.º 86/278/CEE, de 12 de Junho. Este Decreto-Lei veio impôr requisitos de qualidade para as lamas e para os solos, verificáveis através da conformidade das análises requeridas com os valores limite estabelecidos. O diploma define, ainda, um conjunto de restrições à utilização das lamas, bem como deveres de registo de informação por parte dos operadores de gestão de lamas. A intenção deste regime jurídico reside na necessidade de regularizar a utilização agrícola de lamas de depuração, congregando dois objetivos ambientais primordiais: a credibilização da operação de valorização de resíduos e a proteção do ambiente (DL-276/2009).

Outro fim que se pode dar às lamas é a **valorização energética** que tem como objetivo a queima dos componentes orgânicos combustíveis das lamas, de forma a que os produtos resultantes sejam relativamente inertes. Este sistema possui a capacidade de destruição de organismos patogénicos e substâncias perigosas, transformando a lama em cinzas. Estas cinzas poderão ser parte integrante de alguns produtos úteis, nomeadamente materiais de construção (CE, 2001).

Por último o **aterro**, este tipo de fim das lamas só deve ser utilizado quando não é possível proceder à valorização, isto é, quando estão presentes metais pesados ou substâncias tóxicas. Os riscos de contaminação das águas subterrâneas, por lixiviação, e a produção de odores desagradáveis, pela geração de condições anaeróbias e sépticas que exigem uma correta monitorização de todo o processo (CE, 2001).

4. Caracterização do Modelo GPS-X da Hydromantis

Este modelo foi desenvolvido pela empresa Hydromantis com sede no Canadá e consiste num programa de modelação e simulação do comportamento dinâmico dos processos de tratamento de águas residuais domésticas e industriais. O programa foi concebido de acordo com os mais recentes processos de modelação e simulação e utiliza numerosas ferramentas de produtividade que simplificam e flexibilizam a sua aplicação, nomeadamente, no que refere à construção do modelo, simulação e interpretação dos resultados (Carmo, 2013).

O software GPS-X permite processar modelos, quer em estado estacionário, quer em estado dinâmico. O programa fornece dados de entrada muito flexíveis e resultados em quadros, gráficos ou ficheiros que podem ser lidos e trabalhados em outros programas, como no *Microsoft Excel* ou em *MatLab*.

Os relatórios criados pelo simulador do GPS-X são automáticos, e estes são diretamente editados em formato Excel e incluem imagens, gráficos, definições de parâmetros, e uma apresentação muito completa dos resultados da simulação. Os relatórios podem ser de um tipo já previamente definido ou podem ser adaptados de acordo com as orientações do utilizador, em função das suas necessidades (Hydromantis, 2006).

Este software apresenta algumas vantagens nomeadamente, na facilidade de utilização e em particular, no uso das ferramentas do *Influent Advisor* e do *Model Developer* que permitem inovar e desenvolver aplicações específicas, face a objetivos de simulação particulares.

O *Influent Advisor* permite completar a composição das águas residuais afluentes à ETAR quando não se dispõe de uma caracterização completa, mas apenas os valores de alguns dos parâmetros.

O *Model Developer* é um módulo que permite ao utilizador adicionar os seus próprios modelos ou alterar os modelos pré-definidos através de uma aplicação em Microsoft Excel, baseada na representação matricial, utilizando matrizes de Petersen, sem ser necessário introduzir um código específico. Este módulo inclui o cálculo automático dos coeficientes para as diferentes componentes do modelo e é especialmente útil para adicionar componentes industriais aos modelos existentes, por exemplo em caso de ocorrência de toxicidade (Ferreira F. , 2006; Carmo, 2013).

4.1. Esquema do tratamento anaerobio através do modelo GPS-X

Foi realizado um esquema onde retrata a fase da digestão anaeróbia pois só esta fase interessa para elaboração deste trabalho. A figura seguinte representa esse esquema que retrata o primeiro passo deste modelo.

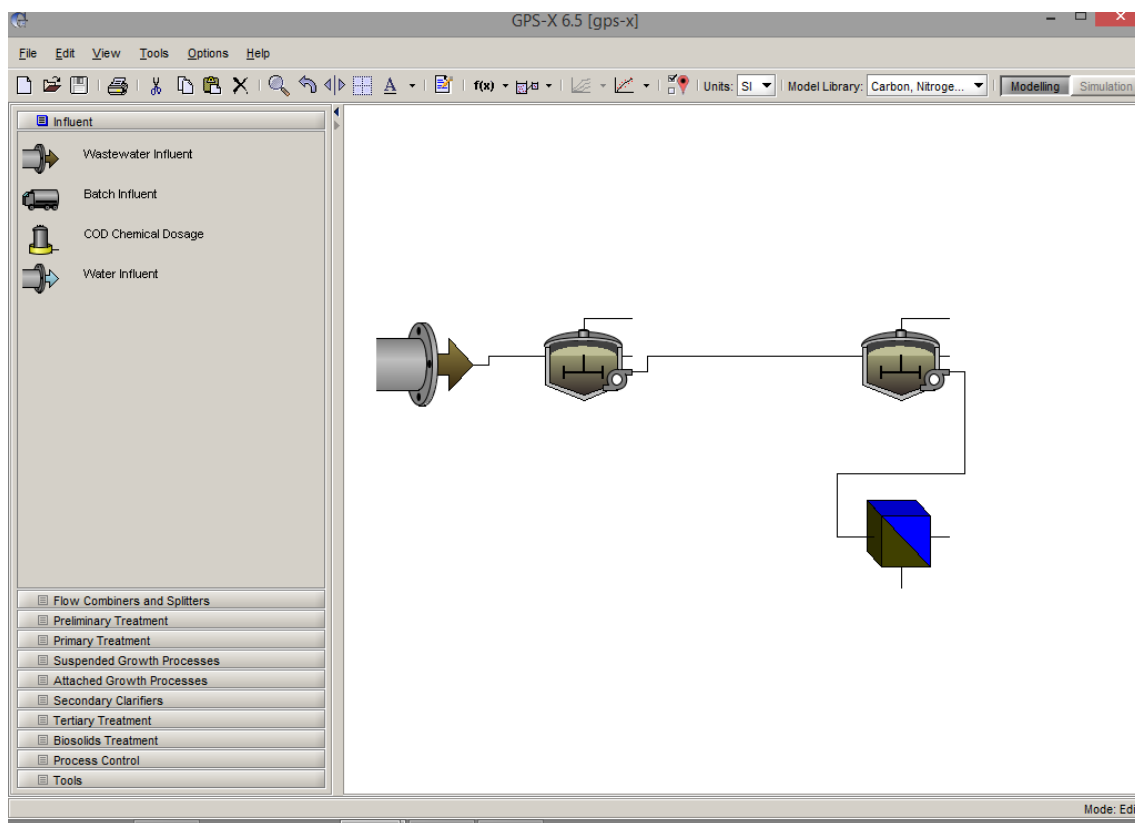


Figura 16- Esquema do tratamento anaeróbio da ETAR do Choupal, desenhado no simulado GPS-X.

Neste trabalho não foi desenvolvido este modelo devido ao tema presente ser bastante ambicioso, ao periodo curto de duração do estágio curricular e visto que, a empresa não disponha de histórico de dados suficiente para a construção deste modelo. Não sendo viável a continuação deste modelo o foco deste relatório ficou-se pelas Redes Neurais Artificiais.

5. Redes Neurais Artificiais

Uma rede neuronal artificial (RNA) é um sistema de processamento de informação que tem algumas semelhanças ao nível do desempenho com as redes neuronais biológicas, este tipo de rede é a principal fonte de inspiração para os modelos neuronais artificiais.

A complexidade das estruturas elementares das redes neuronais biológicas é muito maior do que a dos modelos matemáticos usados nas RNA, demonstra assim as dificuldades encontradas para se tentar imitar o funcionamento do sistema nervoso humano. Sabendo que o sistema nervoso é formado por bilhões de células nervosas, enquanto uma rede neuronal artificial pode possuir dezenas ou máximo milhares de unidades de processamento.

Alguns modelos de redes neuronais artificiais possuem muitos neurónios conectados numa estrutura de peso de conexão e com facilidade de adaptação, proporcionando uma estrutura paralela. Esta estrutura é favorável pois se algum neurónio falhar, os efeitos na rede não são considerados significativos para o desempenho do sistema, surgindo então a tolerância à falha. (Cardon & Müller, 1994)

Segundo (Matos, 2008) é cada vez mais comum a utilização de RNA, têm sido desenvolvidas e aplicadas em várias áreas da ciência de forma a responder às necessidades impostas. O processamento de linguagem e imagem, o reconhecimento de padrões, a classificação de dados, a aproximação de funções, a otimização, o planeamento, a monitorização, o diagnóstico e o controlo são exemplos de campos onde podem ser aplicadas estas redes.

Algumas vantagens em trabalhar com as RNA:

- **Aprendizagem** – uma RNA é treinada com um conjunto de dados que lhe é fornecido, e esta aprende a produzir resultados desejados para esse conjunto, através das intensidades das ligações entre os neurónios;
- **Generalização** – quando é adicionado um novo conjunto de dados, diferente dos exemplos conhecidos, a rede tem a capacidade de produzir resultados para esse conjunto, tendo em conta os exemplos usados anteriormente;
- **Robustez** – mesmo com a falha de alguns neurónios das redes, é possível estas apresentarem um bom desempenho;
- **Paralelismo massivo potencial** – no decurso do processamento da informação, os neurónios “disparam” simultaneamente.

As RNA também tem a capacidade de auto-organização e de processamento temporal, o que faz com que estas sejam uma ferramenta computacional extremamente poderosa e atrativa para a solução de problemas complexos. Estes benefícios permitem que as RNA consigam encontrar boas soluções mesmo com dados sujeitos a ruídos, falhas, informações imprecisas ou corrompidas (Fernandes, 2015).

Devido à complexidade deste processo, tem-se aplicado modelos matemáticos como as RNA para se conseguir entender melhor o processo da digestão e otimizar a produção de biogás, constituindo ainda um sistema de apoio à decisão na gestão do processo anaeróbio (Rocha, 2013).

5.1. Evolução Histórica

Pode-se considerar que a história das Redes Neurais Artificiais começou por volta do ano de 1943 quando McCulloch e Pitts realizaram o primeiro modelo de neurónio artificial baseado na modelagem matemática do neurónio biológico.

Donald Hebb em 1949 escreveu o livro “A organização do comportamento” onde afirmava que a eficiência de uma sinapse variável entre dois neurónios aumenta quando ocorre a ativação repetida de um pelo outro através dessa sinapse (Barreto, 2002).

Ainda citando (Barreto, 2002) em 1958, Frank Rosenblatt cria um novo modelo, o Perceptrão que consistia num filtro linear com a capacidade de adaptação, tratando-se da forma mais simples de uma rede neuronal.

O estudo feito por Widrow e Hoff que consistia no desenvolvimento de um algoritmo de treino baseado no erro quadrático médio (EQM) que implementaram na Adaline (Adaptative Linear Element). Passados dois anos, era proposto por Widrow a Madaline (Multiple-adaline), uma das primeiras redes neurais com vários componentes adaptáveis que tinha a capacidade de aprender (Engel).

A utilização de estruturas com vários elementos foi reforçada em 1963 por Winograd e Cowan, que estes demonstraram que o número elevado de elementos pode representar coletivamente um conceito único, com um aumento correspondente de robustez e paralelismo.

Citando (Barreto, 2002) e referindo (Matos, 2008), na década 60 apesar dos grandes sucessos nesta área, foram feitos estudos matematicamente feitos por Minsky e Papert

onde chegaram à conclusão que havia limites nas operações que perceptrões de camada única podem executar, declararam que não havia razões para assumir que qualquer desses limites pudesse vir a ser ultrapassado na versão multilinear.

Na década de 80, renasce o estudo destas redes com Rumelhart, Hinton e Williams a publicarem o algoritmo de retro-propagação do erro que, aliado à utilização de redes neurais de camadas múltiplas (perceptrões multicamada ou PMC), implementando funções de ativação não lineares, fizeram com que as limitações apresentadas por Minsky e Papert fossem ultrapassadas.

5.2. Analogia com sistemas biológicos

O cérebro humano é o grande processador de informação em que todas as funções e movimentos do organismo estão a cargo dos 10 bilhões de neurónios existentes no sistema nervoso.

Citando (Vieira & Roisenberg), os neurónios estão ligados entre si através de sinapses e juntos formam uma grande rede, a rede neuronal. Dependendo dos impulsos recebidos, um neurónio pode enviar um sinal para outros, através do seu único axónio que liga às dendrites de outros neurónios. Esta conexão é feita por sinapses que transmitem estímulos por todo o corpo através de diferentes concentrações de iões de sódio (Na^+) e potássio (K^+).

As principais características de uma célula neuronal são apresentadas de seguida (Vieira & Roisenberg):

- **Dentrites:** recebem os estímulos transmitidos pelos outros neurónios;
- **Corpo celular do neurónio:** tem como função entender e combinar as informações de outras células neuronais;
- **Axónio:** responsável por transmitir os estímulos para outras células. De seguida apresenta-se a figura 19 da constituição de uma célula neuronal.

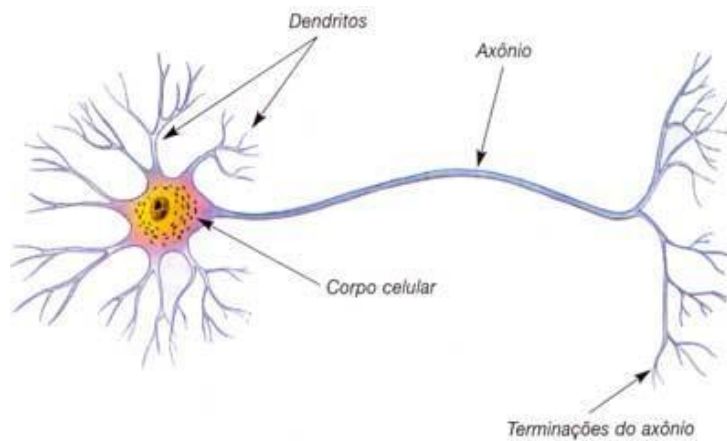


Figura 17- Constituição de uma célula neuronal.

Verificamos que existe uma membrana que envolve cada neurónio que o isola dos outros, ou seja, a comunicação tem que ser feita por via elétrica e por via química através de substâncias neurotransmissores. O impulso é recebido, o neurónio processa a informação, de seguida dispara um segundo impulso que produz uma substância neurotransmissora que se desloca no corpo celular para o axónio. A comunicação entre neurónios é feita na região onde o sinal passa do axónio de um neurónio (pré-sináptico) para a dendrite de outro neurónio (pós-sináptico) sendo designada por sinapse, onde não existe qualquer ligação física (Vieira & Roisenberg; Matos, 2008).

Como já foi referido, McCulloch e Pitts, em 1943, conceberam o seu modelo computacional e lançaram as fundações das RNA pretendiam construir um modelo digital do neurónio biológico.

Assim como o sistema nervoso, a estrutura de uma rede neuronal é composta por unidades conectadas designadas como nós ou neurónios, em que cada um recebe valores de inputs, efetua cálculos relativamente simples com os mesmos e gera um output. Logo, o valor de output de um neurónio é transmitido como um dos inputs de outro neurónio ou como output final de todo o sistema. No seu conjunto, uma RNA é um sistema que recebe inputs, computa-os e devolve um ou mais outputs (Palisade, 2015). A arquitetura típica das redes neuronais apresenta diferentes camadas, nomeadamente, a camada de entrada (inputs), a camada de saída (output) e, pelo menos, uma camada intermédia ou oculta. Mesmo que as RNA possam demonstrar potencialidades interessantes, o sistema nervoso é uma rede de complexidade admirável, incomparavelmente superior às RNA modernas sob todos os pontos de vista. As construções mentais derivam do processamento de informação a vários níveis, desde interações moleculares até circuitos inter-regionais e ao

próprio sistema nervoso central. Tal como os neurónios artificiais são básicos quando comparados com os biológicos, também as RNA são básicas quando comparadas com os circuitos locais encontrados no cérebro (Matos, 2008; Carreira, 2014).

5.3. Princípios de funcionamento e descrição das RNA

Citando (Matos, 2008), uma rede neuronal consiste num sistema de elementos de processamento simples designados por nós ou neurónios que se ligam, armazenando informação de forma implícita na topologia e nos pesos das conexões ou ligações e têm capacidade de se ajustar. Assim sendo, os neurónios são a unidade base das RNA, estando interligados através de conexões ou canais de comunicação, permitindo a passagem de informação. Eles recebem informação de vários inputs, relacionando-os através do peso que estes têm no sistema, somando-os e por fim ditando um resultado que é o output da rede. Este valor é transmitido pela função de transferência, que transforma o valor produzido através da soma ponderada dos inputs que a rede recebe (Silveira, 2009; Brandão, 2014).

Para perceber o funcionamento da rede é necessário conhecer o nó que é a base de uma RNA, demonstrado na próxima figura (Matos, 2008):

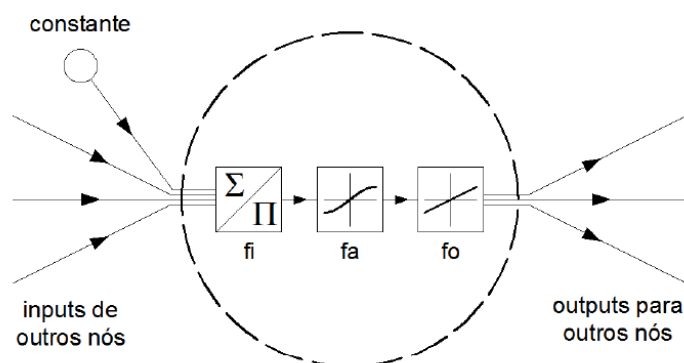


Figura 18- Modelo de um nó de uma RNA (Matos, 2008).

Como se pode ver pela figura o nó recebe ligações de três nós a montante e está ligado a três nós para jusante. De seguida encontram-se as funções de transformação que podem ocorrer no interior do nó (Matos, 2008).

- **Função de Input (fi)** – corresponde ao somatório ou produto dos valores de input num único valor real;

- **Função de ativação (fa)** – correspondência entre o resultado da função de input a um nível de ativação do nó, normalmente varia entre -1 e 1 ou 0 e 1;
- **Função de output (fo)** – destina-se a produzir um sinal a partir da ativação do nó mas nem sempre é utilizada.

Como já referido existem vários tipos de RNA, podendo ser diferenciadas pela estrutura, no género de computações realizadas dentro dos neurónios até mesmo nos algoritmos de treino (Silveira, 2009). Sendo as redes mais comuns aquelas constituídas por neurónios ordenados em camadas, o que faz com que cada neurónio receba como inputs apenas os outputs dos neurónios das camadas anteriores, ou seja, não existem ligações entre neurónios da mesma camada.

As redes neuronais conseguem modelar funções de, praticamente, qualquer complexidade, estando o número de camadas e de nós em cada camada, associado à complexidade da função (Silveira, 2009; Matos, 2008).

O perceptrão e o perceptrão multicamada podem ser considerados os modelos de redes neuronais mais usuais ao nível da arquitetura.

Perceptrão

É um tipo de rede constituída apenas por uma camada, sendo que neste caso, geralmente, a rede é constituída por um único neurónio e um polarizador, com pesos ajustáveis. Esta é considerada a rede neuronal mais simples com capacidade de classificar padrões linearmente separáveis. O algoritmo de treino deste tipo de rede foi o primeiro modelo de treino supervisionado, ou seja, para conseguir efetuar o treino é necessário deter a classificação desejada para cada um dos vetores que fazem parte dos dados (Matos, 2008; Rocha, 2013).

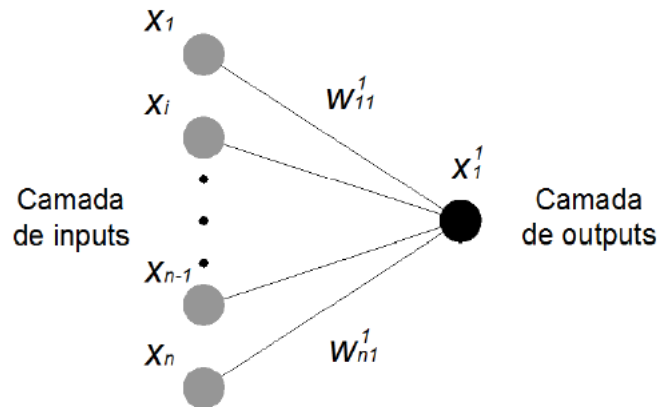


Figura 19- Representação esquemática de um perceptron (Matos, 2008).

Referindo (Matos, 2008) e citando (Brandão, 2014) a computação do valor introduzido no nó de output é feita através do produto entre o valor de cada input (X_i) e o peso da respetiva ligação ao nó de input (W_{i1}^1), resultando em y , que representa o número real que é transformado em X_1^1 , derivado do somatório de todos os produtos.

O teorema da convergência do perceptron foi criado por Rosenblatt, consistia se os padrões utilizados no treino do perceptron pertencem a duas classes linearmente separáveis, o algoritmo do perceptron converge e posiciona a superfície de decisão no hiperplano entre essas classes. Posteriormente, evidenciou-se que aumentando o número de nós, o número de classes que o perceptron permite separar também sofre um aumento, mesmo tendo constituído uma importante inovação (Rocha, 2013; Brandão, 2014).

Perceptron multicamada (PMC)

Referindo (Matos, 2008) e citando (Carreira, 2014) este modelo é dos mais conhecidos na literatura e pode também ser chamado por *Feed-Forward Nets* ou *Perceptron Nets*, consiste em possuir uma camada de input, uma camada de output e, pelo menos, uma camada intermédia ou oculta, em que cada nó está ligado a todos os nós da camada seguinte.

Na próxima figura, apresenta-se um exemplo de uma rede multicamada com uma camada intermédia.

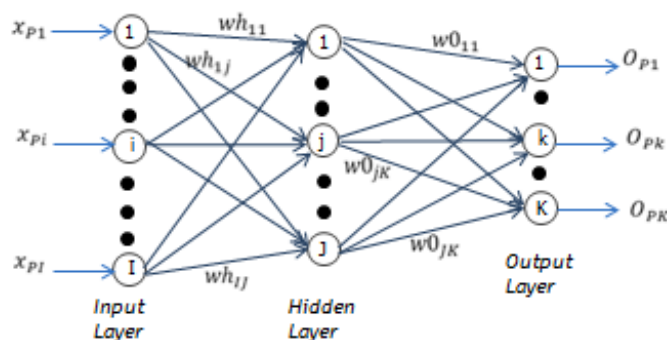


Figura 19- Esquema das ligações e elementos presentes num PMC de uma camada oculta (Matos, 2008).

Estes modelos são úteis por constituírem aproximadores universais, tendo o potencial para aproximar qualquer função contínua com nível de erro abaixo de qualquer superior a 0, tal como é estabelecido pelo teorema do aproximador universal. Na prática é demonstrado que tal resultado não é muitas vezes aplicável, visto que a quantidade de nós e dados necessários para o atingir aumenta relativamente com a complexidade da função e com a diminuição do erro (Matos, 2008; Brandão, 2014).

Ainda referindo (Matos, 2008) e (Brandão, 2014) em relação a soluções de aproximação tradicionais, bem como polinomiais ou trigonométricas, as PMC têm vantagens pois a degradação da qualidade da aproximação com o aumento de dimensões do problema é bastante menor. Tendo como desvantagem a sua complexidade uma vez que as suas operações que ocorrem no interior da rede são muito complexas pelo que o treino, deve seguir procedimentos que evitem tanto quanto possível problemas na solução final.

5.4. Tipos de Redes Neurais Artificiais

As redes neurais podem ser classificadas segundo:

- **Arquitetura**- podem ser de redes planares ou de camadas;
- **Processo de aprendizagem**- o treino pode ser feito com ou sem supervisão;
- **Dinâmica**- se existe transferência de informação sequencial ou simultânea.

Neste trabalho foi utilizado o software NeuralTools da Palisade, esta ferramenta utiliza redes neurais Multicamada *Feed-Foward* (MLF) ou Perceptrão Multicamada (MLP), redes de Regressão Generalizada (GRNN) e redes Probabilísticas (PNN). As redes MLF são utilizadas quer para previsões numéricas quer para previsões categóricas, já as outras

duas dividem-se, as GRNN servem para fazer previsão numérica e as PNN para previsões categóricas (Rocha, 2013; Carreira, 2014).

Neste estudo em concreto pretende-se uma previsão numérica, sendo assim, foram utilizadas redes MLF e GRNN.

Redes MLF e PNN/GRNN

Referindo (Rocha, 2013) e (Fernandes, 2015), uma rede neuronal multicamada é constituída por camadas alinhadas de neurónios ou nós, é capaz de efetuar aproximações de funções complexas, podendo modelar relações complexas entre variáveis independentes e uma variável dependente. Relativamente à arquitetura, estas redes apresentam uma camada de entrada (input), uma camada de saída (output) e, pelo menos, uma camada intermédia ou oculta.

Os neurónios de uma camada estão ligados apenas aos neurónios da camada seguinte, não havendo ligações entre neurónios da mesma camada e a comunicação é efetuada num único sentido. O número de nós na camada de entrada corresponde ao número de variáveis que serão usadas para alimentar a rede neuronal e os nós assumem o valor dos dados de entrada (Rocha, 2013; Fernandes, 2015).

As redes PNN e GRNN funcionam da mesma maneira, podendo diferenciar-se pelo tipo de problema a que podem ser aplicadas. Sendo assim, se a variável dependente for categórica utiliza-se a rede PNN, caso a variável dependente for numérica, recorre-se a uma rede GRNN (Fernandes, 2015).

Em relação às redes GRNN segundo (Rocha, 2013) e citando (Brandão, 2014), são estruturadas da seguinte maneira: uma camada de inputs, cada nó representa uma variável de entrada; duas camadas ocultas, a primeira camada oculta (camada padrão) contém um nó para cada caso de treino; a segunda camada oculta (camada de soma) é constituída por nós de numeradores e denominadores e o seu tamanho depende de diversos fatores relacionados com os dados de treino. Cada neurónio da camada de padrões calcula a sua distância em relação ao caso apresentado. Os valores passados aos nós de numeradores e denominadores são funções da distância e do valor dependente. Os nós da camada de soma efetuam a soma de seus inputs, enquanto o nó de output os divide para gerar a previsão, já sendo a fase final deste tipo de rede. Quando um caso é apresentado à rede, esta calcula o valor dependente previsto utilizando os valores dependentes de cada caso

de treino, sendo que os casos de treino mais próximos contribuem de forma mais significativa ao valor do output.

O treino de uma rede GRNN baseia-se em otimizar os fatores de suavização para minimizar o erro no conjunto de treino. A medida de erro, usada durante o treino, para avaliar diferentes conjuntos de fatores de suavização é o erro médio quadrático (EMQ).

5.5. Fases de desenvolvimento de uma RNA

No desenvolvimento de uma RNA temos de considerar 4 fases distintas: A preparação de dados; o treino; o teste e a previsão da rede (Palisade, 2015). A seguir apresenta-se a descrição das 4 fases.

Fase de Preparação dos dados

Esta fase fundamentalmente consiste em dois passos muito importantes, numa boa compreensão do processo a modelar e na preparação dos dados de input, ou seja, compreender muito bem o processo para que a escolha e a preparação dos dados a definir seja feita da melhor maneira, para evitar erros ou utilizar dados menos fiáveis, assim deve-se escolher as variáveis de entrada que realmente interessam no processo e também identificar as variáveis de saída, aquelas que se pretende fazer a previsão.

Referindo (Lingireddy S. and Brion G. M., 2005) citado por (Carreira, 2014), esta pode ser considerada a etapa mais longa e importante, devido à recolha extensa de dados e à escolha das variáveis, pois vão ter um grande peso neste processo.

Como já foi dito, em relação à base de dados é muito importante a quantidade e qualidade, dois fatores fulcrais. Dado que, quanto maior for o número de dados beneficia a compreensão por parte da rede, melhorando e aumentando o desempenho da rede bem como a capacidade de generalização. Em relação à qualidade dos dados, é importante manter a sua garantia para que os resultados obtidos sejam credíveis, aumentando a fiabilidade dos dados escolhidos (Carreira, 2014).

Definindo assim a base de dados, procede-se à realização de três conjuntos nomeadamente, conjunto de treino, conjunto de teste e conjunto de validação. O conjunto de treino é aquele que reúne a maior quantidade de casos, estes que são conhecidos e que serão utilizados no processo de aprendizagem da rede. O conjunto de teste é utilizado para avaliar a capacidade de generalização da rede, ou seja, o valor que se pretende prever

é conhecido e será comparado com aquele que é obtido. E por último, o conjunto de validação, que serve para confirmar resultados considerados razoáveis na fase de teste, garantido a capacidade de generalização da rede, este conjunto pode ou não ser feito (Carreira, 2014).

Fase de Treino da RNA

É nesta fase que é feito o processo de aprendizagem de forma a ajustar os parâmetros de computação da rede (Palisade, 2015), os valores do conjunto de dados escolhidos para o treino são conhecidos e pretende-se que através dos dados de entrada, sejam obtidos resultados de output o mais próximo possível dos valores reais.

O treino é o processo de aprendizagem que ajusta os parâmetros de computação da rede, sendo o objetivo, fazer com que a RNA retorne valores aproximadamente corretos face a dados inputs. O processo é orientado pelos dados fornecidos para treino e pelo algoritmo de treino, este último seleciona vários conjuntos de parâmetros e avalia cada um aplicando a RNA a cada caso de treino e vendo a qualidade da resposta da rede. O algoritmo seleciona novos conjuntos de parâmetros face aos resultados das anteriores tentativas. Os dados que são utilizados para o processo de treino não são todos disponibilizados, pois na fase de teste e ou numa fase de validação serão necessários (Brandão, 2014).

Fase de Teste da RNA

De acordo com (Palisade, 2015), nesta fase verifica-se a capacidade de prever valores de output conhecidos, mas estes não faziam parte do conjunto de treino.

Terminado o teste, o desempenho da rede é medido estatisticamente para verificar a percentagem de respostas conhecidas que foram previstas corretamente. O erro médio quadrático é determinado nesta fase e este é utilizado para escolher a melhor rede, ou seja, a rede que tiver menor valor de erro é considerada a rede com maior garantia de sucesso na previsão de novos valores desconhecidos (Brandão, 2014).

Fase de Previsão da RNA

Nesta fase a rede já foi treinada e testada e está pronta para ser utilizada com valores de output desconhecidos, sendo este o principal objetivo das RNA (Palisade, 2015).

A previsão comparativamente com os processos anteriores é mais simples, o treino pode ser feito várias vezes, com configurações de rede diferentes, até chegar a uma rede que

na fase de teste mostre melhores resultados. Para tal as redes têm de adquirir capacidade de generalização necessária para que seja assim possível escolher a melhor e efetuar previsões com um certo grau de confiança (Carreira, 2014).

6. Apresentação do Caso de Estudo

6.1. Enquadramento do Estudo

A otimização da produção de biogás, no processo de DA, é um processo essencial para a rentabilidade de exploração de uma ETAR.

Devido a esta necessidade de otimizar o máximo possível da produção de biogás tem-se apostado na utilização das RNA que permitem fazer o estudo do processo da digestão anaeróbia através da análise e tratamento dos seus dados com a finalidade de permitir uma melhor gestão desse mesmo processo. O objetivo do estudo passa por relacionar as variáveis do processo com a produção de biogás, de forma a otimizá-lo e ajudar na tomada de decisões no processo de produção. Espera-se, assim, que estas ferramentas matemáticas sejam capazes de contribuir para a compreensão do processo, assumindo-se como um sistema de apoio da gestão da digestão anaeróbia (Carreira, 2014).

Este tipo de estudo tem vindo a ser muito realizado nas diferentes ETAR do grupo ADP, e como tem havido bons resultados ao utilizar as RNA para conseguir uma melhor produção de biogás no processo de DA a ETAR do Choupal não quis perder esta oportunidade.

6.2. Metodologia Proposta

Para o desenvolvimento de uma parte do trabalho utilizaram-se os programas Microsoft Excel e NeuralTools, do The Decision Tools Suite, da Palisade.

6.2.1. Apresentação do NeuralTools

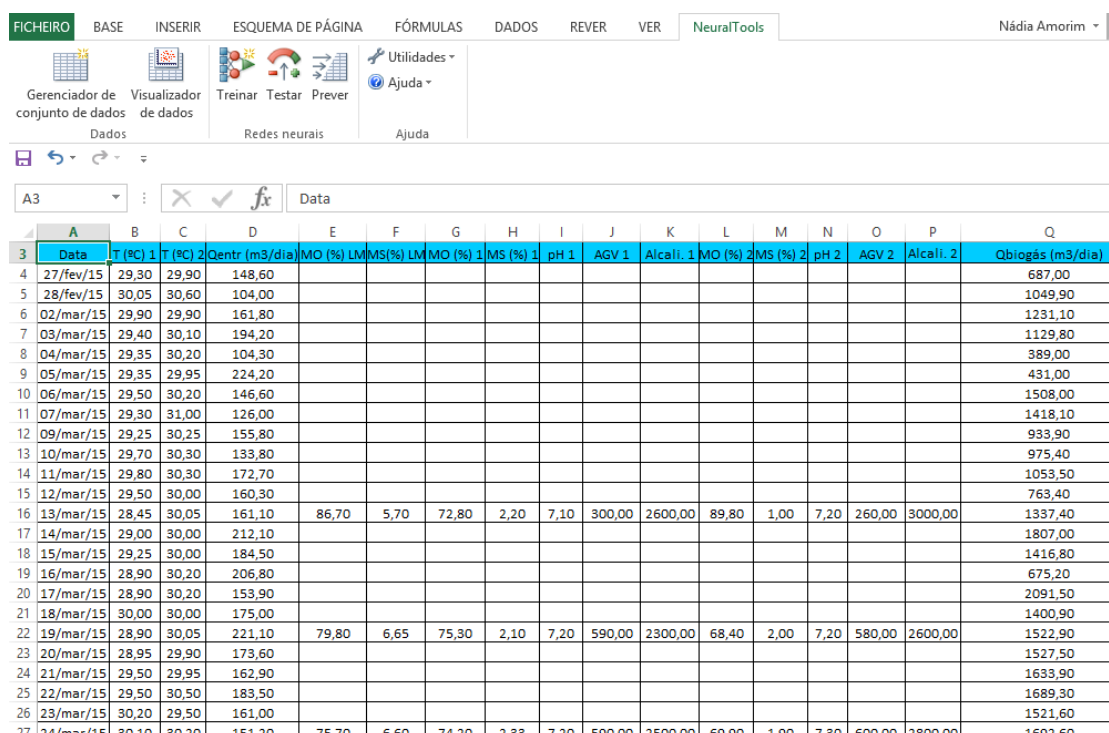
De acordo com (Palisade, 2015) e citado por (Carreira, 2014), esta ferramenta informática é usada no Microsoft Excel, com objetivo a efetuar previsões através de recursos avançados de redes neuronais, este possui um gerenciador de dados avançado e algoritmos bem desenvolvidos, sendo fácil de usar. Como se trabalha diretamente no Excel uma ferramenta utilizada com mais frequência ajudou na execução dos procedimentos do NeuralTools, pois permite utilizar dados já inseridos manualmente ou utilizando as várias funcionalidades do próprio programa. Este software consegue resolver dois tipos de problemas nomeadamente de classificação ou numéricos e tal permite a utilização quer de variáveis categóricas ou variáveis numéricas, podendo estas ser dependentes ou independentes, criando relações lineares ou não-lineares.

Como já foi dito anteriormente a base de dados selecionada tem que ser grande, conseguindo assim passar pelas diversas fases da RNA, o desenvolvimento destas serão explicadas de seguida.

6.2.2. Preparação de dados

A preparação dos dados para criar e trabalhar na rede neuronal é um passo muito importante para a continuação do trabalho.

O software NeuralTools trabalha sobre as folhas de cálculo do Excel, em que as colunas representam as variáveis intervenientes e as linhas correspondem, uma a uma, a casos com diferentes valores para as variáveis escolhidas. Cada caso possui um conjunto de valores de variáveis independentes, inputs, e um valor para a variável dependente, output, este último podendo ser conhecido ou não. Na figura 20 é apresentada um exemplo de uma base de dados para ser utilizada no NeuralTools.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	Data	T (°C) 1	T (°C) 2	Qentr (m3/dia)	MO (%)	LMMS (%)	LM MO (%)	1 MS (%)	1 pH	AGV 1	Alcali. 1	MO (%) 2	MS (%) 2	pH 2	AGV 2	Alcali. 2	Obiogás (m3/dia)
4	27/fev/15	29,30	29,90	148,60													687,00
5	28/fev/15	30,05	30,60	104,00													1049,90
6	02/mar/15	29,90	29,90	161,80													1231,10
7	03/mar/15	29,40	30,10	194,20													1129,80
8	04/mar/15	29,35	30,20	104,30													389,00
9	05/mar/15	29,35	29,95	224,20													431,00
10	06/mar/15	29,50	30,20	146,60													1508,00
11	07/mar/15	29,30	31,00	126,00													1418,10
12	09/mar/15	29,25	30,25	155,80													933,90
13	10/mar/15	29,70	30,30	133,80													975,40
14	11/mar/15	29,80	30,30	172,70													1053,50
15	12/mar/15	29,50	30,00	160,30													763,40
16	13/mar/15	28,45	30,05	161,10	86,70	5,70	72,80	2,20	7,10	300,00	2600,00	89,80	1,00	7,20	260,00	3000,00	1337,40
17	14/mar/15	29,00	30,00	212,10													1807,00
18	15/mar/15	29,25	30,00	184,50													1416,80
19	16/mar/15	28,90	30,20	206,80													675,20
20	17/mar/15	28,90	30,20	153,90													2091,50
21	18/mar/15	30,00	30,00	175,00													1400,90
22	19/mar/15	28,90	30,05	221,10	79,80	6,65	75,30	2,10	7,20	590,00	2300,00	68,40	2,00	7,20	580,00	2600,00	1522,90
23	20/mar/15	28,95	29,90	173,60													1527,50
24	21/mar/15	29,50	29,95	162,90													1633,90
25	22/mar/15	29,50	30,50	183,50													1689,30
26	23/mar/15	30,20	29,50	161,00													1521,60
27	24/mar/15	30,10	30,20	151,20	75,70	6,60	74,20	2,33	7,20	590,00	2500,00	69,90	1,90	7,30	600,00	2800,00	1692,60

Figura 20- Exemplo de uma base de dados para utilizar com o NeuralTools.

O seguinte comando “Gerenciador de conjuntos de dados”, este permite definir qual o conjunto de dados a ser utilizado. Ao clicar nesse comando aparece a caixa de diálogo apresentada na figura abaixo, que além de permitir definir o conjunto de dados também permite editar e excluir as variáveis que se pretende utilizar nas várias fases das redes neuronais artificiais.

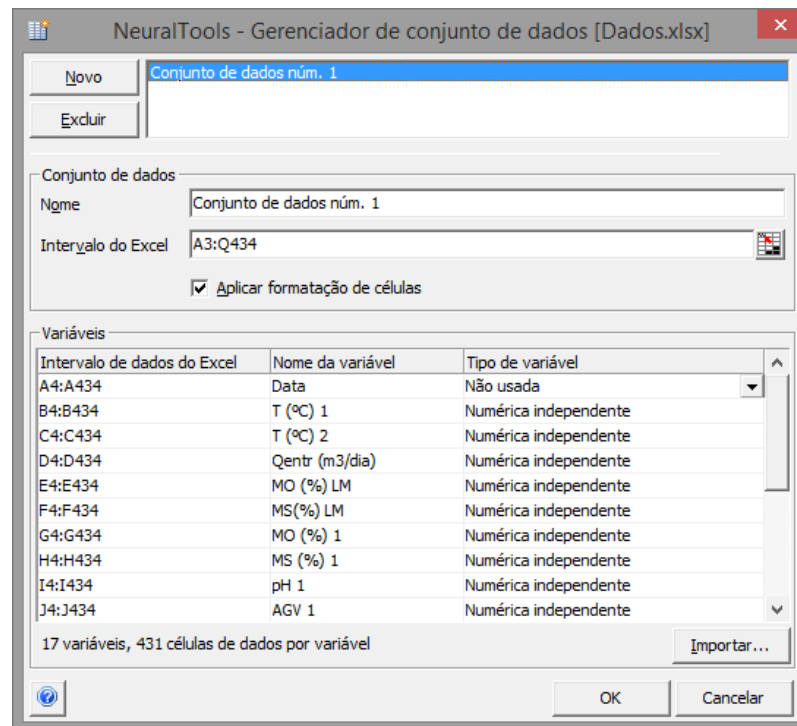


Figura 21- Caixa de diálogo do gestor do conjunto de dados.

A caixa de diálogo apresenta várias opções, tais como:

- **Novo e Excluir** – adiciona um novo conjunto de dados ou exclui um conjunto já existente;
- **Nome** – permite especificar um nome para o conjunto de dados criado;
- **Intervalo do Excel** – especifica o intervalo do Excel correspondente ao conjunto de dados. Se forem selecionados vários intervalos de células no conjunto de dados, este item será antecedido pelo rótulo Múltiplos;
- **Múltiplos** – permite especificar intervalos de células individuais que formam o conjunto de dados, com múltiplos intervalos de células;
- **Aplicar formatação de células** – acrescenta uma grade e cores que identificam os conjuntos de dados;
- **Importar** – permite copiar, neste conjunto de dados, os tipos de variáveis de um outro conjunto de dados ou de uma rede neuronal.

Na secção de “Variáveis” são apresentadas as variáveis disponíveis na base de dados para serem escolhidas como variáveis de entrada do conjunto de dados, havendo informação acerca do intervalo de dados do Excel a que corresponde essa variável, o nome e o tipo de variável.

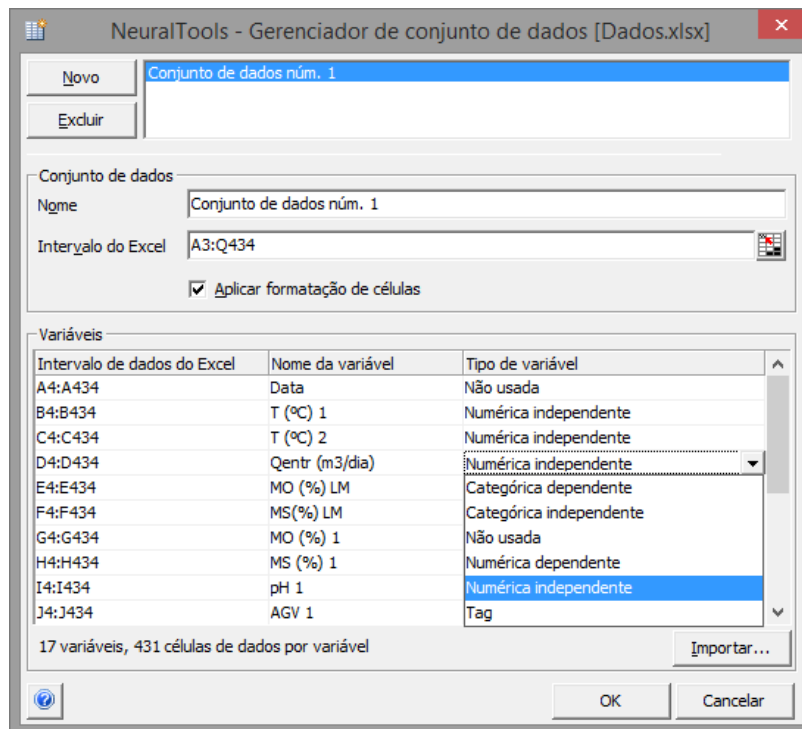


Figura 22- Definição do tipo de variável.

Existem seis tipos de variáveis, como:

- **Categorica dependente** - é uma variável dependente ou de output cujos valores possíveis são tomados de um conjunto de possíveis categorias;
- **Categorica independente** - é uma variável independente ou de input cujos valores possíveis são tomados de um conjunto de possíveis categorias;
- **Numérica dependente** - é uma variável dependente ou de output cujos valores possíveis são numéricos;
- **Numérica Independente** – é uma variável independente ou de input cujos valores possíveis são numéricos;
- **Tag** - é uma variável que assume possíveis valores de “treino”, “teste” ou “previsão”. Esse tipo de variável é escolhido pelo utilizador para identificar casos a serem usados especificamente para treino, teste e previsão;
- **Não usada** – é uma variável da base de dados criada que não será usada na rede neuronal.

Ao fim de ter o conjunto de dados definido e organizar as variáveis, a folha de cálculo do Excel fica como se pode ver na próxima figura.

FICHEIRO

BASE

INSERIR

ESQUEMA DE PÁGINA

FÓRMULAS

DADOS

REVER

VER

NeuralTools

Nádia Amorim

Gerenciador de conjunto de dados

Visualizador de dados

Treinar

Testar

Prever

Utilidades ▾

Ajuda ▾

Dados

Redes neurais

Ajuda

F2

✕

✓

fx

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

Data

T (°C)

1 T (°C)

2 Qentr (m3/dia)

MO (%)

LMMS (%)

LM MO (%)

1 MS (%)

1 pH

AGV 1

Alcali. 1

MO (%)

2 MS (%)

2 pH

AGV 2

Alcali. 2

Consumo de biogás (m3/dia)

25/fev/15

28,70

30,00

21,20

289,10

26/fev/15

29,20

29,90

82,20

572,40

27/fev/15

29,30

29,90

148,60

687,00

28/fev/15

30,05

30,60

104,00

1049,90

01/mar/15

30,00

31,00

87,40

758,60

02/mar/15

29,90

29,90

161,80

1231,10

03/mar/15

29,40

30,10

194,20

1129,80

04/mar/15

29,35

30,20

104,30

389,00

05/mar/15

29,35

29,95

224,20

431,00

06/mar/15

29,50

30,20

146,60

1508,00

07/mar/15

29,30

31,00

126,00

1418,10

08/mar/15

29,90

30,60

43,90

1117,20

09/mar/15

29,25

30,25

155,80

933,90

10/mar/15

29,70

30,30

133,80

975,40

11/mar/15

29,80

30,30

172,70

1053,50

12/mar/15

29,50

30,00

160,30

763,40

13/mar/15

28,45

30,05

161,10

86,70

5,70

72,80

2,20

7,10

300,00

2600,00

89,80

1,00

7,20

260,00

3000,00

1337,40

14/mar/15

29,00

30,00

212,10

1807,00

15/mar/15

29,25

30,00

184,50

1416,80

16/mar/15

28,90

30,20

206,80

675,20

17/mar/15

28,90

30,20

153,90

2091,50

18/mar/15

30,00

30,00

175,00

1400,90

19/mar/15

28,90

30,05

221,10

79,80

6,65

75,30

2,10

7,20

590,00

2300,00

68,40

2,00

7,20

580,00

2600,00

1522,90

20/mar/15

28,95

29,90

173,60

1527,50

Figura 23- Conjunto de dados definidos a ser utilizados.

Ao terminar esta fase de criar e definir o conjunto de dados, está tudo preparado para se seguir para as próximas fases do processo da rede neuronal.

6.2.3. Treino e Teste da Rede Neuronal

Na fase de treino, os valores de output são conhecidos, sendo assim treinada a rede neuronal. Já na fase de teste há a previsão de valores output através da rede neuronal treinada, medindo assim o seu desempenho. O NeuralTools permite que estas fases sejam feitas em simultâneo, ou seja, do conjunto de dados criado, parte é utilizada para o treino da rede neuronal e parte para o teste, sendo essa percentagem definida pelo utilizador.

Na figura seguinte é apresentada uma caixa de diálogo do “Treinamento” onde se pode encontrar os comandos “treinar”, “ Configuração da rede” bem como o “Tempo de execução” onde se pode definir todas as condições para que ocorra o treino da rede.

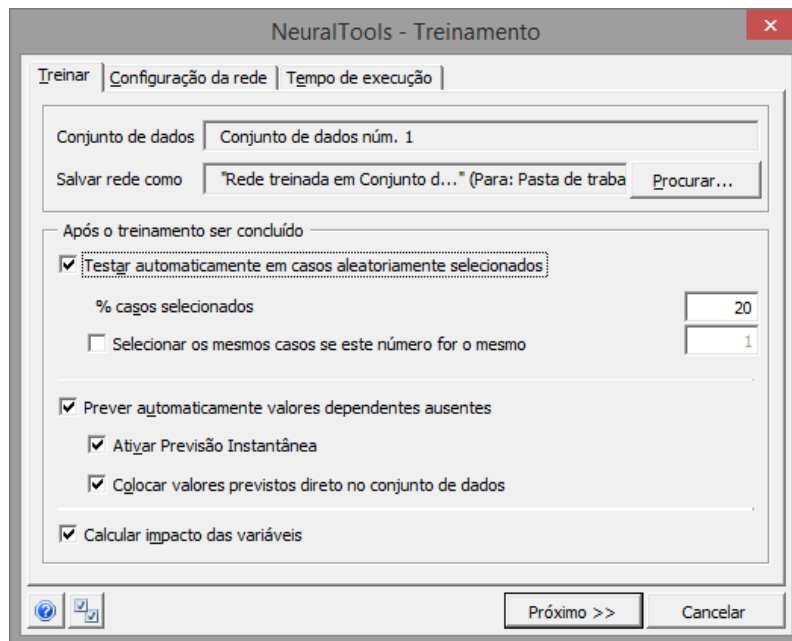


Figura 24- Caixa de diálogo do treino.

No separador “treinar” da caixa de diálogo, define-se qual o conjunto de dados que será utilizado no treino, bem como o local onde a rede será guardada no Excel. É aqui que também existe a opção de efetuar o teste seguido do treino, tornando as duas fases numa só, através da opção “Testar automaticamente em casos aleatoriamente selecionados”, em que é definida uma percentagem de casos que são “excluídos” do treino, para poderem ser utilizados apenas no teste da rede neuronal. Outra das opções também presente é “Prever automaticamente valores dependentes ausentes”, em que esses dados são também usados para a fase da previsão.

Quando é ativada esta opção, dispõe-se de mais duas escolhas, a “Previsão Instantânea”, em que é permitido alterar o valor das variáveis independentes o que altera instantaneamente o valor da previsão, e a “Colocação de valores previstos diretamente no conjunto de dados”. Ainda nesta caixa de diálogo, é possível escolher a opção “Calcular o impacto das variáveis” em que o NeuralTools calcula o impacto que cada variável independente ou de input teve no processo de treino, por outras palavras, mede a sensibilidade das previsões da rede a mudanças nas variáveis referidas.

6.2.3.1. Configuração da rede

Neste separador é escolhido o tipo de rede neuronal a ser utilizado, dispondo, além da escolha específica de uma rede, também da opção da procura da melhor rede, como se pode ver na figura seguinte.

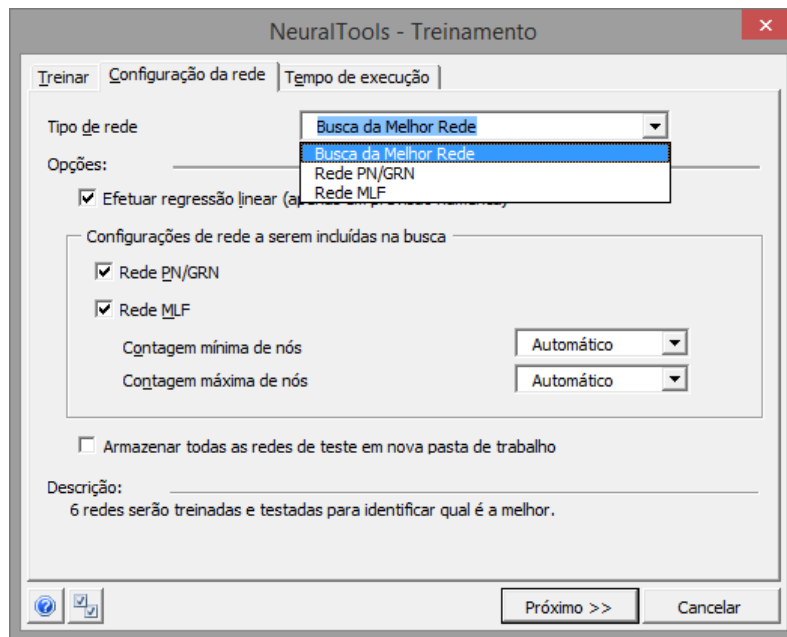


Figura 25- Caixa de diálogo da Configuração da rede.

Como se pode verificar pela figura tem-se três possibilidades de escolha para o treino da rede e para a previsão de classificação/categoria existem dois tipos de redes disponíveis: Redes Neurais Probabilísticas (PNN) e redes Multi-Layer Feedforward (MLF). A previsão numérica pode ser realizada usando-se redes MLF assim como Redes Neurais de Regressão Generalizada (GRNN), que são semelhantes às redes PNN. O NeuralTools possibilita ainda um recurso de busca da Melhor Rede, quando esta opção é selecionada, o NeuralTools treina e testa uma variedade de configurações de redes neuronais até gerar a que produz as melhores previsões para os seus dados. É aconselhável usar sempre a opção busca da melhor rede.

6.2.3.2. Tempo de execução

Em relação a este comando existem três critérios de paragem: duração, número de tentativas e andamento. A “Duração” especifica o tempo de limite que se pretende para o treino. Quando se opta por procurar a melhor rede, cada configuração de rede treinada durará o tempo que é aqui especificado. A opção “Tentativas” especifica o número de tentativas que o NeuralTools deve executar para cada rede. Por último o “Andamento” trata da percentagem mínima de mudança de erro em função do tempo, em que o NeuralTools parará se não puder melhorar a estatística de erros pelo menos na percentagem especificada dentro do intervalo de tempo definido.

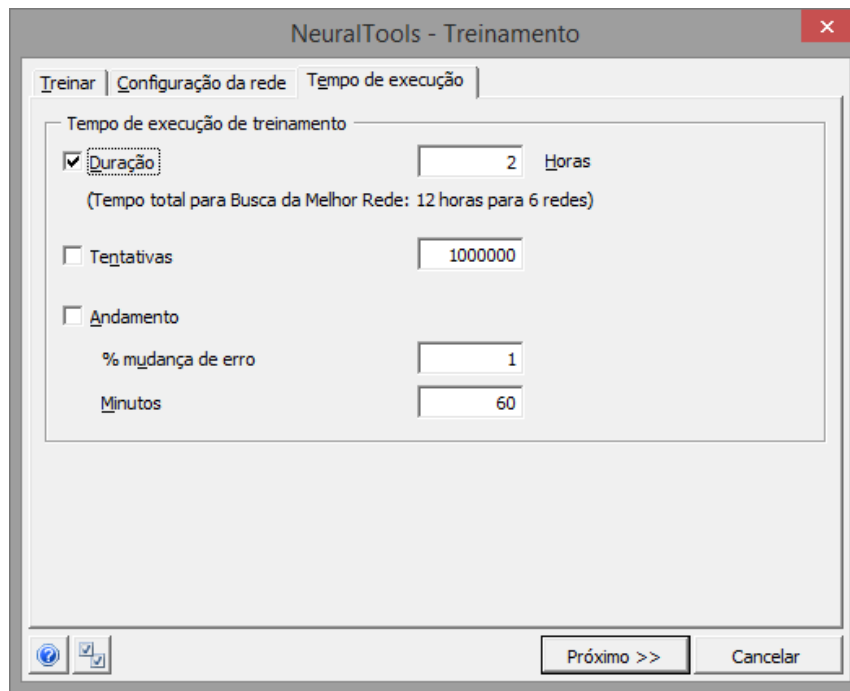


Figura 26- Caixa de diálogo do Tempo de execução.

Não se sabe ao certo quanto tempo demora uma rede neuronal a ser treinada, portanto é preferível aceitar as condições padrão do software.

A fim de definir as condições do algoritmo de treino, passa-se para a última fase anterior ao treino. Na caixa de diálogo abaixo apresentada observa-se a visão geral do treino, apresenta-se um resumo com a informação acerca das configurações de treino e do conjunto de dados, bem como as variáveis intervenientes.

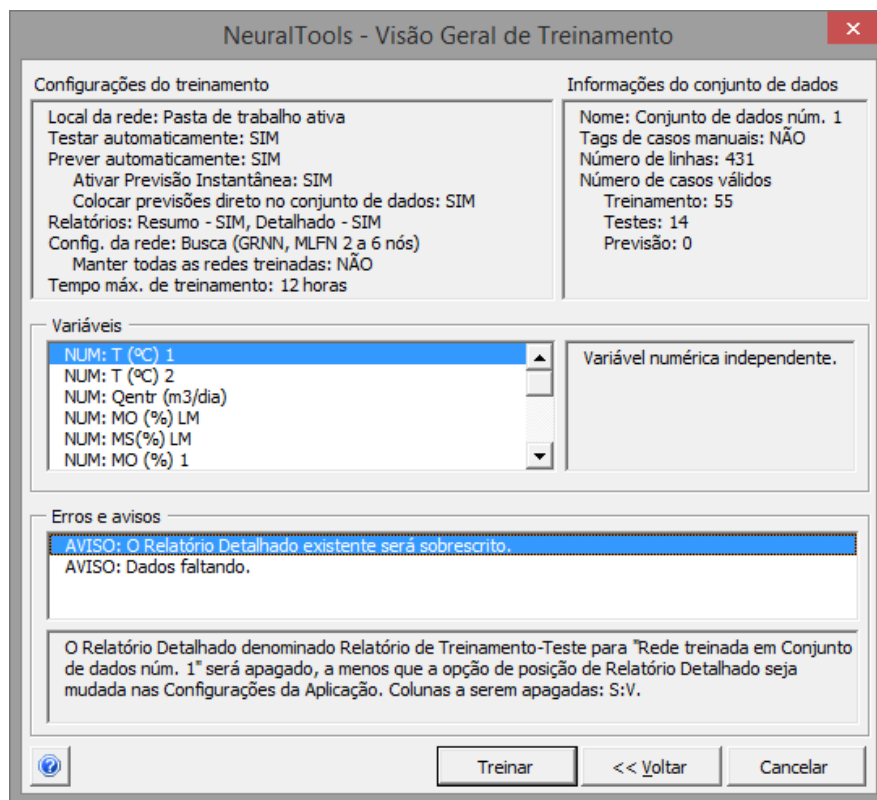


Figura 27- Caixa de diálogo da Visão geral de treino.

É importante ter em atenção a caixa de diálogo na secção “Erros e avisos” pois caso exista alguma anomalia com os dados ou com alguma das opções escolhidas na configuração o NeuralTools deteta e é transmitido nesta mesma caixa de diálogo, evitando assim tempo perdido a treinar a rede pois esta é a fase mais demorada.

6.2.3.3. Processo de Treino

Ao iniciar o treino, é possível observá-lo em tempo real, isto é, na figura seguinte onde se encontra caixa de diálogo do andamento do treino surgem gráficos em que é possível observar os erros à medida que decorre o treino de cada configuração. Nesta fase as mudanças nos erros observadas nos gráficos nada permite concluir acerca da qualidade das previsões que a rede fará com relação aos casos que não foram usados no treino. O treino é interrompido quando se atinge algum dos critérios de paragem definidos.

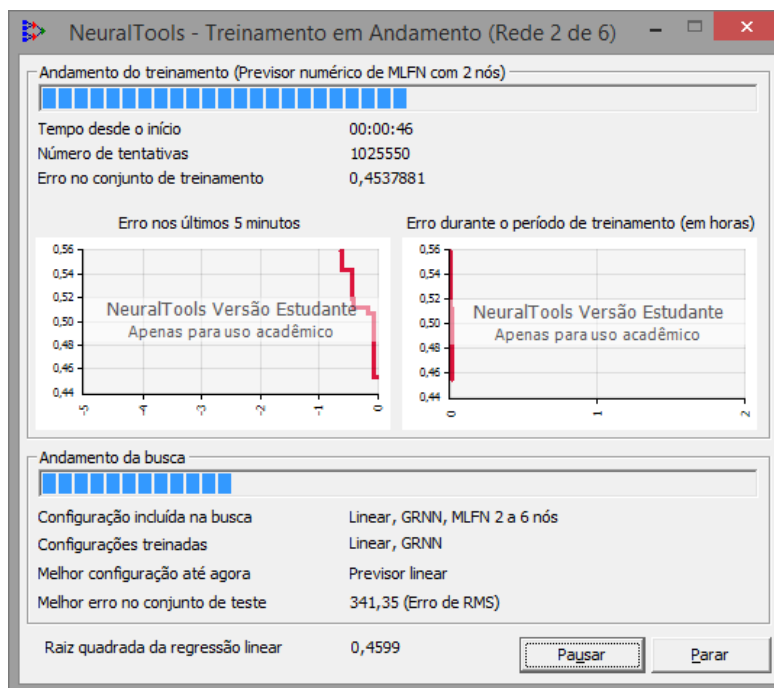


Figura 28- Caixa de diálogo do Treinamento em Andamento.

6.2.3.4. Relatório de Treino

Esta é a última fase do treino da rede, onde são criados automaticamente relatórios detalhados que permitem analisar a rede em termos de desempenho, como também fornecem informação acerca da rede em termos de configuração, treino e teste. Relativamente ao treino e teste, obtém-se conhecimento acerca do número de casos em causa, percentagem de previsões inadequadas, erro médio quadrático e absoluto e desvio padrão do erro absoluto. O cálculo do impacto das variáveis é determinado desde que se tenha ativado essa função na caixa de diálogo do treino.

Como exemplo, será apresentado de seguida um relatório de Treino e Teste automático de uma rede neuronal, com a devida explicação dos principais parâmetros estatísticos.

Tabela 3- Tabelas geradas no relatório de treino e teste automático do NeuralTools

NeuralTools: Treinamento e Teste Automático de rede neural

Executado por: Nádia Amorim

Data: 5 de agosto de 2016 05:18:04

Conjunto de dados: Conjunto de dados núm. 1

Rede: Rede treinada em Conjunto de dados núm. 1 (62)

Resumo	
Informações da rede	
Nome	Rede treinada em Conjunto de dados núm. 1 (62)
Configuração incluída na busca	GRNN, MLFN 2 a 6 nós
Melhor configuração	Previsor numérico de MLFN (2 nós)
Local	Esta pasta de trabalho
Variáveis categóricas independentes	0
Variáveis numéricas independentes	4 (T (°C) 1, Qentr (m3/dia), MO (%) LM, MS(%) LM)
Variável dependente	Var. numérica (Qbiogás (m3/dia))
Treinamento	
Número de casos	50
Tempo de treinamento	02:00:00
Número de tentativas	89160108
Motivo da parada	Parada automática
% previsões inadequadas (15% tolerância)	54,0000%
Erro médio quadrático	252,07
Erro médio absoluto	194,98
Desvio padrão de erro absoluto	159,76
Teste	
Número de casos	13
% previsões inadequadas (15% tolerância)	61,5385%
Erro médio quadrático	358,60
Erro médio absoluto	288,89
Desvio padrão de erro absoluto	212,45
Conjunto de dados	
Nome	Conjunto de dados núm. 1
Número de linhas	398
Tags de casos manuais	NÃO
Análise de impacto de variáveis	
T (°C) 1	34,0218%
Qentr (m3/dia)	33,6438%
MO (%) LM	27,9477%
MS(%) LM	4,3867%

Busca da Melhor Rede			
	Erro de RMS	Tempo de treinamento	Motivo da parada do treinamento
GRNN	418,42	00:00:00	Parada automática
MLFN 2 nós	358,60	02:00:00	Parada automática
MLFN 3 nós	396,37	02:00:00	Parada automática
MLFN 4 nós	420,67	02:00:00	Parada automática
MLFN 5 nós	561,53	02:00:00	Parada automática
MLFN 6 nós	419,19	02:00:00	Parada automática

- **% previsões inadequadas** – uma previsão é considerada “inválida” se estiver fora da margem definida em torno do valor efetivo;
- **Erro médio quadrático** – medida do desvio das previsões em relação ao valor efetivo. É calculado fazendo a raiz quadrada do desvio quadrático médio.
- **Erro médio absoluto** – desvio médio das previsões em relação aos valores efetivos.
- **Impacto das variáveis** – exibe o impacto relativo das variáveis independentes ou de input nas respostas previstas.

Neste relatório encontram-se gráficos de dispersão mostrando assim as relações existentes entre os valores efetivos, previstos e residuais. Um valor residual é diferença entre o valor efetivo e o valor previsto.

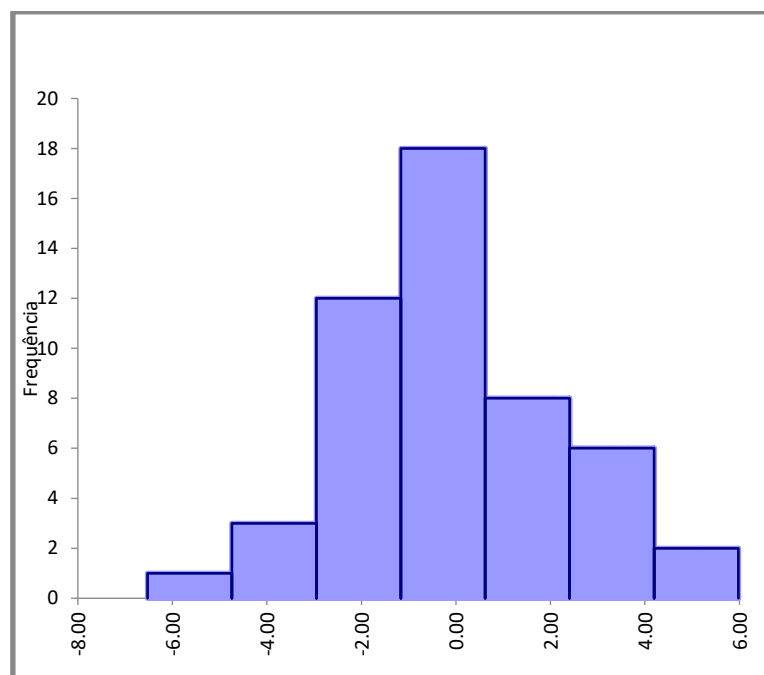


Figura 29- Gráfico das diferenças residuais, Treinamento.

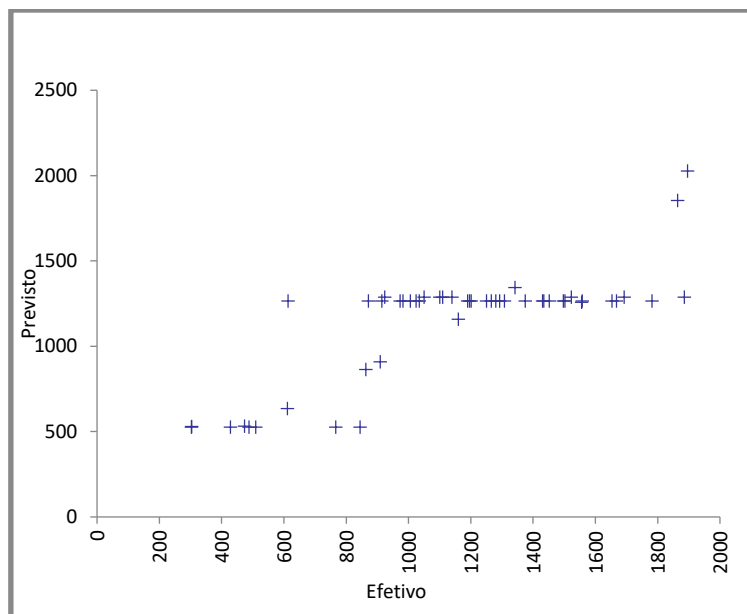


Figura 30- Gráfico relativo aos dados previstos comparados com os dados efetivos, Treinamento.

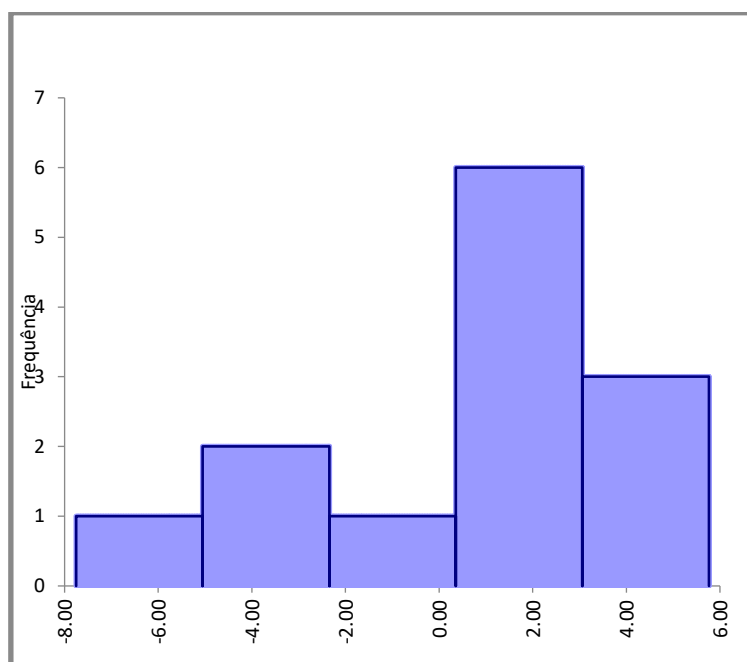


Figura 31- Gráfico referente às diferenças residuais, Teste.

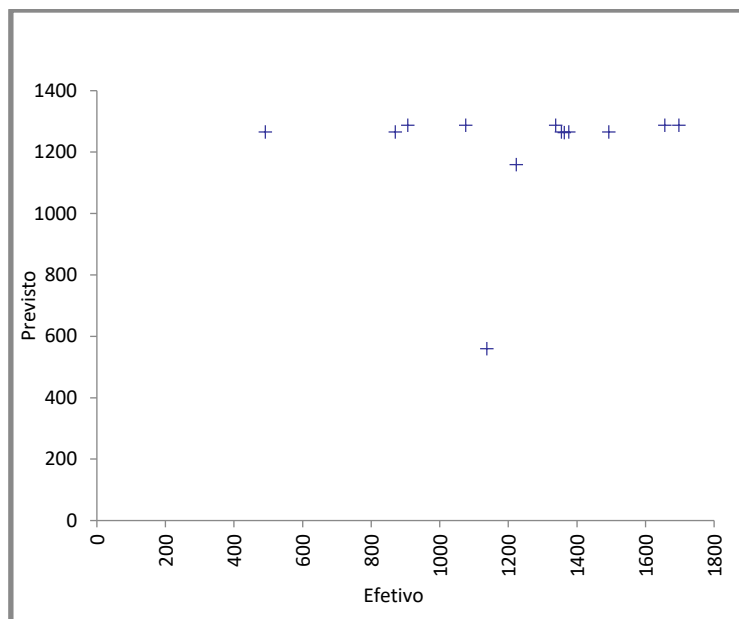


Figura 32- Gráfico dos dados previstos comparando com os dados efetivos, Teste.

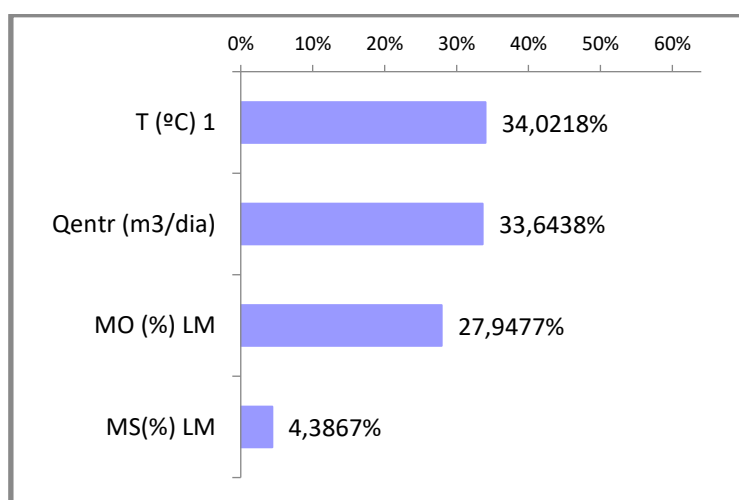


Figura 33- Gráfico do impacto das variáveis.

6.2.4. Previsão

Ao fim destas fases a rede neuronal está pronta para atingir o seu objetivo, prever valores de output desconhecidos em novos casos. O NeuralTools disponibiliza dois métodos de previsão, são eles:

- 1) Método de previsão de valores desconhecidos da variável de output de um conjunto de dados, para isso utiliza-se o comando existente no NeuralTools;
- 2) Previsão Instantânea, onde os valores das variáveis de input de determinados casos podem ser alterados e a partir daí são previstos valores de output com as novas condições.

Esta previsão permite modificar os dados de *input* e observar as alterações que essas mudanças podem provocar. Quando é ativada este tipo de previsão o NeuralTools acrescenta automaticamente uma fórmula de Excel na célula que contém o valor previsto. Portanto, se os valores das variáveis independentes de um caso específico forem mudados, o valor previsto será automaticamente recalculado, por essa fórmula. Ao usar a Previsão Instantânea, basta digitar os dados referentes aos novos casos diretamente no Excel e gerar uma nova previsão automaticamente, sem passar pela caixa de diálogo Previsão.

6.3. Desenvolvimento do Estudo

6.3.1. Introdução

Logo no início do estágio, a primeira etapa foi adquirir o software necessário para a realização deste estudo, o NeuralTools, esta aquisição foi feita a partir da empresa AdCL.

Antes de iniciar o trabalho, foi feito um conhecimento geral sobre a ETAR do Choupal, nomeadamente das instalações, os operadores, e do processo de funcionamento da digestão anaeróbia da ETAR. Neste ponto foi muito importante a ajuda do co-orientador da AdCL que esteve sempre pronto para responder às dúvidas que iam surgindo.

Sendo que nesta ETAR os digestores anaeróbios funcionam em simultâneo em regime em série, ou seja, o estudo das RNA foi feito com os dois digestores (Digestor 1 e Digestor 2), o conjunto de dados a tratar estão diretamente relacionados com os dois digestores anaeróbios, e claramente que a produção de biógas é uma variável presente, para se conseguir atingir o objectivo deste trabalho.

6.3.2. Preparação dos dados

A recolha dos dados necessários para a elaboração deste trabalho foram feitos através de um software (NAVIA) implementado pela empresa, este permite o armazenamento dos dados recolhidos ao longo do tempo, conseguindo assim, uma grande base de dados para o treino das RNA. Segundo (Lingireddy S. and Brion G. M., 2005) e citado por (Rocha, 2013), este processo de preparação de dados é essencial pois a quantidade e a qualidade destes são muito importantes para o sucesso do treino da RNA, parte-se do princípio que, quanto maior for a base de dados, melhor será o treino e a capacidade de previsão da rede, logo espera-se que seja mais fácil perceber as relações existentes entre dados.

São recolhidos um variadíssimo leque de dados, nomeadamente de recolha de lamas e de efluentes. Existem amostras diárias, semanais, quinzenais, mensais, trimestrais, etc., os operadores é que se ocupam por esta tarefa de fazer as recolhas. A sua periodicidade depende do processo implicado para a sua obtenção, pois existem dados que são recolhidos e obtidos logo no local de amostragem, e outros dados necessitam de tratamento laboratorial, ficando sempre dependentes da disponibilidade do laboratório. Os resultados das amostras são carregadas no software, permitindo consultar o histórico dos dados recolhidos dos últimos anos.

Na realização deste trabalho, foram utilizados os dados diários e semanais relativos aos meses do ano de 2015 até Junho de 2016. Em anexo são apresentadas tabelas com alguns desses dados, pois existe uma grande quantidade de dados e seria muito extenso colocar aqui as tabelas completas, assim sendo, optou-se por só colocar uma amostra dos dados que foram utilizados neste trabalho.

Verificamos que as variáveis que foram recolhidas e utilizadas como entradas na modelação das redes neuronais de previsão da produção de biogás foram: a percentagem de matéria seca (% MS_{LM}) e orgânica (%MO_{LM}) das lamas mistas, que entra no digestor I, o caudal de lama que entra para o digestor I ($Q_{entr.(m^3/dia)}$) que também é o mesmo para o digestor II, o caudal de biogás ($Q_{biogás(m^3/dia)}$), a percentagem de matéria seca (%MS), a percentagem de matéria orgânica (%MO), a temperatura (T°C), o pH, a alcalinidade, e os ácidos gordos voláteis (AGV em mgCH₃COOH/l) de cada um dos digestores. De seguida, serão descritas estas mesmas variáveis:

- **%MS**- representa a percentagem de matéria sólida presente nas lamas.
- **%MO**- representa o conteúdo em matéria orgânica da lama.
- **Q_{entr.}**- Representa o volume de lamas que entram no digestor ao longo de um dia, este dependente da quantidade e características de efluente que é encaminhado para a ETAR.
- **pH** – representa o valor do pH das lamas completamente misturadas dentro do digestor.
- **Alcalinidade** – representa a capacidade de resistência a alterações de pH causadas pela adição de ácidos (acidificação). A alcalinidade resulta da presença de hidróxidos, carbonatos e bicarbonatos e é expressa em termos de carbonato de cálcio. A concentração da alcalinidade em águas residuais é importante em

sistemas de remoção biológica de nutrientes, como é o caso da DA. A estabilidade do digestor é reforçada por uma elevada concentração da alcalinidade.

- **Ácidos Gordos Voláteis (AGV)** – são os principais produtos intermédios formados durante o processo de DA. A presença de altas concentrações de ácidos voláteis provoca a diminuição da concentração da alcalinidade e uma descida do pH, podendo tornar o meio tóxico para algumas bactérias presentes no processo de digestão.
- **Temperatura**- refere-se à temperatura das lamas dentro do digestor. É usado o valor médio de três medições de temperatura efetuadas ao longo do dia. Como já foi referido no presente trabalho o controlo e a manutenção de uma temperatura constante no digestor é crucial para o bom funcionamento do processo anaeróbio, uma vez que flutuações pontuais e/ou frequentes da temperatura podem provocar efeitos severos. Neste caso de estudo o digestor opera em regime mesófilo num intervalo de temperaturas entre 28°C a 39°C. A operação fora de um intervalo ótimo de temperatura (acima ou baixo) pode provocar inibição do processo de digestão.
- **Q_{biogás}** – este parâmetro representa o caudal diário de produção do biogás que é produzido pela DA, apresentado em m³/dia. Esta variável foi calculada partir da soma do caudal de consumo com o caudal excesso.

Sendo assim, estas variáveis representam os inputs, e como já referido o objetivo da modelação do processo de DA através de RNA é a otimização da produção de biogás, logo, caudal de produção de biogás foi definido como output.

Como já foi referido anteriormente, os dados recolhidos para fazer o treino e teste das redes neuronais a serem modeladas, referem-se aos digestores que trabalham em série e estão compreendidos entre Fevereiro de 2015 até Junho de 2016. Assim sendo, foi possível elaborar a base de dados que consta no Anexo I. Durante a elaboração desta base de dados foram analisados os valores das variáveis, caso a caso, para tentar eliminar possíveis erros de medição ou digitação de valores.

6.3.3. Treino e Teste

Iniciou-se esta fase após de se ter feito um estudo exaustivo dos dados e ter-se definido um bom conjunto de dados para a modelação de redes neuronais. Apesar na parte inicial deste trabalho estas etapas serem descritas separadamente, o NeuralTools permite que

estas aconteçam em simultâneo do ponto de vista do utilizador, facilitando o seu trabalho. Este executa em primeiro lugar o treino e depois o teste sem que haja necessidade de intervenção do utilizador. O software tem a capacidade de dividir o conjunto de dados fornecido em conjunto de treino e conjunto de teste, mas dispõe a possibilidade de se poder definir a percentagem dos dados disponíveis que se pretende que intervenham na fase de teste. Assim, depois de serem definidas todas as configurações da rede, o NeuralTools escolhe aleatoriamente os casos pertencentes a cada fase.

Podem ser treinadas e testadas várias redes, mas no fim só uma é escolhida para a previsão dos resultados finais, sendo assim, a escolha da melhor rede é feita de acordo com as previsões inadequadas na fase de teste e com o erro médio quadrático (EMQ) que possui. A melhor rede é aquela que apresenta um valor menor em cada um destes dois parâmetros. Para entender melhor o funcionamento do software, foram criados inicialmente conjuntos de dados aleatórios com o objetivo explorar e compreender todas as funcionalidades da fase de treino e teste de RNA.

Foram realizados três conjuntos de redes, para este processo de treino e teste, cada um desses conjuntos tem redes referentes à base de dados existente.

No primeiro conjunto começou-se por modelar uma rede neuronal usando todas as variáveis de entrada, com o objetivo de ver qual era o impacto que cada variável tinha no processo, de forma a poder definir qual a sua importância. Sendo depois possível criar novos conjuntos de dados, desprezando assim as variáveis com menor impacto, pois de acordo com (Lingireddy S. and Brion G. M., 2005) citando (Rocha, 2013) é uma das maneiras para serem feitos conjuntos de redes de sucesso, logo se seguida, verificar o efeito da redução do número de variáveis na rede, por último remodelar a rede com o reduzido número de variáveis experimentado diferentes combinações até se obter o melhor resultado.

O NeuralTools permite treinar e testar numerosas configurações de redes até encontrar aquela que possui melhor desempenho para o conjunto de dados existentes, as redes foram treinadas no modo em busca da melhor rede. Como o que se pretende trata-se de uma previsão numérica, o NeuralTools dispõe de dois tipos de rede: redes Multi-Layer Feedforward (MLF) ou Redes Neurais de Regressão Generalizada (GRNN). A busca da melhor rede, por definição, são treinadas redes GRNN e MLF de 2 a 6 nós numa camada oculta, havendo uma comparação dos resultados com uma regressão linear, mas

estas configurações podem ser alteradas pelo utilizador escolhendo se quer que seja feita regressão linear, quais dos dois tipos de redes devem ser incluídas na procura e, no caso das redes MLF, existe a opção de escolher o número mínimo e máximo de nós que a camada oculta pode possuir, variando entre 2 e 100. Considerando a principal vantagem das redes GRNN reside na sua rapidez de processamento, quase instantânea, quando comparadas com as MLF, que são muito mais lentas, podendo demorar muitos minutos, ou até horas, para encontrar a melhor solução.

De seguida serão apresentados os três conjuntos de redes.

6.3.3.1. Primeiro conjunto de redes

Como já foi referido anteriormente para este primeiro conjunto de redes começou-se por elaborar uma rede com todas as variáveis de entrada possíveis, para entender o comportamento da rede e visualizar quais as variáveis que tinham maior impacto. Para os seguintes conjuntos de rede foram realizados com o critério de desprezar as variáveis que tinham menor impacto, e de uma maneira aleatória foram feitos vários conjuntos de redes para conseguir encontrar a melhor rede, considerando o EMQ menor.

Ao fim de realizar enúmeros conjuntos e analisar os relatórios detalhados que o NeuralTools realiza, pode-se dizer que a melhor rede encontrada foi o conjunto designado de “RN 11”, constituída pelas variáveis de entrada para o Digestor 1 e pelas variáveis desse mesmo digestor, ou seja, % MO₁, T(°C)₁, % MS_{LM}, % MO_{LM}, % MS₁, pH₁, Qentr_(m3/dia), Alcal₁, AGV₁. Esta rede tem um erro EMQ de 126,24 e foi treinada com uma configuração MLFN (3 nós) e possui cerca de 18% de previsões inadequadas, as restantes redes obtiveram um EMQ superior e inclusive maior valor de previsões inadequadas.

No Anexo II apresenta-se a tabela dos conjuntos de redes elaborados para este primeiro conjunto, bem como, um exemplo dos relatórios gerados pelo NeuralTools para rede.

6.3.3.2. Segundo conjunto de redes

A partir da melhor rede do conjunto anterior foram feitos testes desprezando as variáveis que apresentavam menos impacto, e de uma maneira aleatória e tentando diminuir o número de variáveis foi-se escolhendo variáveis que se acharia que tivessem uma maior influencia na rede. Assim, foram feitos conjuntos de redes tentado sempre encontrar a rede que tivesse menor erro médio quadrático.

Ao analisar os conjuntos de redes realizados, podem ser consultados no anexo IV, pode-se dizer que a rede que conseguiu ter melhor erro foi a “RN2”, com uma configuração GRNN e constituída pelas seguintes variáveis: % MS_{LM}, T(°C)₁, Alkali.₁, pH₁, % MO₂, Qentr_(m3/dia), % MS₁, % MO₁, AGV₁, % MO_{LM}, tendo um EMQ de 145,74 e com 24 % de previsões inadequadas.

Comparando com o primeiro conjunto de redes nota-se que foram utilizadas as mesmas variáveis, tendo-se acrescentado a matéria orgânica do digestor 2, possivelmente devido a este factor houve um ligeiro aumento do erro. Mas, de uma maneira geral, conseguiu-se alcançar conjuntos de redes com EMQ razoáveis, nomeadamente as redes “RN10” e “RN11” com valores de 151,31 e 159,55 respectivamente, e nestes conjuntos já se conseguiu reduzir o número de variáveis. Cada rede foi constituída por seis variáveis sendo elas diretamente relacionadas com o digestor 1 e que entram para este, conseguindo inclusive que a “RN11” tivesse uma percentagem de previsão inadequada mais baixa do que a “RN2”, tendo como valor de 22 %.

6.3.3.3. Terceiro conjunto de redes

Neste conjunto reduziu-se as variáveis uma vez que a rede é beneficiada se for treinada com um número reduzido de variáveis, sendo este um dos fatores mais importante para conseguir uma boa rede, como explicado anteriormente.

Posto isto, a rede conseguida de forma mais positiva para o resultado neste conjunto é a rede designada por “RN9” composta por cinco variáveis nomeadamente, Qentr_(m3/dia), % MS_{LM}, % MS₁, T(°C)₁, % MO_{LM}. Esta rede conseguiu atingir um valor de EMQ de 146,20 e uma percentagem de previsão inadequada de 26%. Neste mesmo conjunto também se conseguiu outras duas redes com erros médios quadráticos aceitáveis, a rede “RN6” e a rede “RN7”, constituídas por seis variáveis são elas, % MO₁, % MO_{LM}, % MS₁, pH₁, T(°C)₁ e o Qentr_(m3/dia) para a primeira rede, a diferença para a segunda rede houve uma troca de variáveis, ou seja, trocou-se % MO₁ e pH₁ pelas variáveis AGV₁ e % MS_{LM} as restantes mantiveram-se igual. Estas redes foram obtidas partir da configuração MLFN (3 nós), atingiram um valor de EMQ de 152,24 e 147,0 respectivamente, sendo que a rede “RN7” conseguiu um menor valor de percentagem de previsão inadequada, de 16%.

No fim de analisar os conjuntos realizados e visto que as melhores redes são constituídas por variáveis diretamente relacionadas com o digestor 1, pensou-se em testar uma rede onde só estivessem presentes as seguintes variáveis, $T(^{\circ}\text{C})_1$, $Q_{\text{entr}(m3/\text{dia})}$, % MO_{LM} , % MS_{LM} . Analisando o relatório detalhado gerado pelo NeuralTools esta rede obteve melhor configuração MLFN (2 nós) e um EMQ de 252,07 e 54% de previsão inadequadas.

Conclui-se também que as variáveis com maior impacto na modelação do processo anaeróbio são a % MO_1 , a % MO_{LM} , a $T(^{\circ}\text{C})_1$, o $Q_{\text{entr}(m3/\text{dia})}$. Os resultados obtidos por estas redes podem ser consultados no Anexo V, representados numa tabela simplificada para facilitar a análise das redes. Nessa tabela, assim como nas anteriores, podem-se consultar os valores do EMQ e a percentagem de más previsões em cada fase, as variáveis de entrada para cada rede e o impacto de cada uma na modelação da rede, entre outras informações.

Algumas redes apresentam valores muito altos de EMQ, o que pode dever-se a alguns fatores. A escolha das variáveis não ser a mais correta, conjuntos como foram feitos de uma maneira aleatória a escolha de algumas variáveis podem ter um contributo negativo para a rede, valores mal introduzidos durante a recolha de amostras, a base de dados ser curta, ou seja, o intervalo de tempo devia ser maior pois como já foi referido as RNA tem melhores resultados tendo um leque maior de amostras. Como se pode comprovar esta base de dados tem muitos dias que não foram feitas recolhas, isso também pode prejudicar a rede neuronal.

6.3.4. Validação

Foi criado um conjunto de dados de validação para conseguir comparar os valores previstos pela rede e os valores reais, este conjunto de dados é composto pelos valores de funcionamento do digestor, obtidos nos meses de Maio e Junho de 2016. Este conjunto tem como finalidade avaliar a capacidade de generalização da rede.

Ao comparar os resultados previstos com os valores reais neste conjunto, pode-se verificar que as redes sentem alguma dificuldade em se aproximar, mas ainda assim o comportamento é bastante semelhante, ocorrendo uma certa tendência para os valores previstos acompanharem os valores reais. Para ajudar a visualizar esta comparação de

dados, foi realizado um gráfico que pode ser consultado no anexo VII, bem como, a base de dados para a elaboração deste gráfico.

Pode-se dizer que esta dificuldade sentida dos valores previstos com os valores reais se deve a diversos fatores, como por exemplo, o conjunto de dados para validação da rede ser muito pequeno, deveria-se aumentar assim a amostra de dados para conseguir avaliar o comportamento da previsão de uma forma mais completa, erros induzidos durante a recolha de amostras ou valores recolhidos que podem não ser representativas da realidade do processo, ou até mesmo a escolha das variáveis não ser a correta podendo haver a necessidade de mudar certas variáveis que possam ter um contributo mais positivo no processo de digestão.

6.3.5. Previsão

Sendo esta a última etapa do processo de modelação da RNA e tem como finalidade verificar a capacidade de generalização e o desempenho da melhor RNA escolhida na fase de treino, neste caso a rede que foi escolhida foi a última rede treinada constituída por quatro variáveis.

Nesta fase criou-se um novo conjunto de dados do qual as redes não tiveram conhecimento na fase de treino e teste, para depois efetuar previsões de biogás produzido associado às variáveis de input das redes, sem conhecer os valores de output.

Este ponto baseou-se em analisar o comportamento da melhor rede na previsão de produção de biogás, realizando gráficos tridimensionais fazendo variar num eixo a percentagem de matéria orgânica (% MO) e no outro o tempo de retenção hidráulica (TRH). Este conjunto de dados não foi baseado em dados reais, foram definidos valores dentro da gama de funcionamento dos digestores. Os valores das variáveis foram organizados em tabelas para posteriormente serem realizados os gráficos para ser feita uma leitura mais fácil. Estes gráficos podem ser consultados no anexo VIII que ajudam na interpretação dos dados de previsão.

Conclui-se então que a rede treinada conseguiu aproximar-se do comportamento do processo da digestão anaeróbia, chegando a ter valores na gama de funcionamento ótimo.

7. Considerações finais

7.1. Conclusões

No desenvolvimento do estudo constatou-se a importância das estações de tratamento de águas residuais para prevenir a poluição ambiental e evitar as consequências que os resíduos possam ter para a saúde pública. Contudo, a ETAR's têm grandes consumos energéticos associados às operações de tratamento, para colmatar estas dificuldades e no intuito de aumentar a rentabilidade de exploração da ETAR recorre-se a soluções económicas, ecológicas e ambientais. Para tal recorre-se à otimização do processo da digestão anaeróbia cujo o produto final é o biogás, uma fonte de energia renovável, com baixo custo associado.

As RNA são uma ferramenta ainda em desenvolvimento mas nos estudos e projetos já realizados na área da modelação do processo da digestão anaeróbia tem havido bons resultados. Consta-se que a aplicação de modelos de simulação permite prever a resposta dinâmica das várias unidades de tratamento de uma ETAR a diversas perturbações, possibilitando a implementação de estratégias que garantam um melhor e maior desempenho da mesma. Devido ao seu sistema de processamento de informação, as RNA possuem algumas características de desempenho muito semelhantes com as redes neuronais biológicas, estas são a grande fonte de inspiração para os modelos neuronais artificiais. Por este tipo de modelação ser muito recente, muito não conhecem que esta ferramenta pode ser útil em variadíssimas áreas, como na ciência, na economia, bem como na saúde, entre outras.

Para desenvolver a modelação do processo da digestão anaeróbia utilizando as RNA, procedeu-se à aquisição do software NeuralTools da Palisade, que pode ser trabalhado no Microsoft Excel. Este envolve quatro fases para elaborar uma RNA, a preparação de dados, o treino, o teste e a previsão da rede. Ao fim de elaborar estas fases descritas ao longo deste trabalho, obteve-se redes do tipo MLFN e GRNN que conseguiram um comportamento interessante, encontrando redes com um bom desempenho na modelação. Ao analisar os resultados obtidos, foi possível encontrar um conjunto de variáveis com maior influência neste processo da DA.

A melhor rede conseguida foi treinada com as variáveis relacionadas com o digestor 1, ou seja, % MO_1 , $T(^{\circ}C)_1$, % MS_{LM} , % MO_{LM} , % MS_1 , pH_1 , $Q_{entr(m3/dia)}$, $Alcal.1$, AGV_1 .

Esta rede obteve um erro EMQ de 126,24 e foi treinada com uma configuração MLFN (3 nós) e possui cerca de 18% de previsões inadequadas.

Pode-se considerar que as variáveis com maior impacto na modelação do processo anaeróbio são a % MO_1 , a % MO_{LM} , a $T(^{\circ}C)_1$, o $Q_{entr(m3/dia)}$.

Visto que as melhores redes treinadas tinham as variáveis relacionadas com o digestor 1, foi treinada uma rede constituída por quatro variáveis, $T(^{\circ}C)_1$, $Q_{entr(m3/dia)}$, % MO_{LM} , % MS_{LM} . Ao analisar o relatório gerado pelo software obteve a melhor configuração MLFN (2 nós), um EMQ de 252,07 e 54% de previsão inadequadas.

Certas redes apresentam valores muitos altos de EMQ, devido a alguns fatores, como a escolha das variáveis não ser a correta, erros mal introduzidos durante a recolha de amostras, a base de dados ser curta, ou seja, o intervalo de tempo devia ser maior pois como já foi referido as RNA tem melhores resultados tendo um leque maior de amostras, e aumentar a periodicidade da recolha das amostras iria beneficiar a rede neuronal.

Em relação aos dados utilizados para validação da rede, nota-se que os valores de previsão tem alguma dificuldade em se aproximar com os valores reais, mas ocorrendo uma certa tendência para os valores previstos acompanharem os valores reais, como se pode verificar pelo gráfico que se encontra no anexo VII. Apesar desta dificuldade de não haver uma aproximação de valores, devido a erros já mencionados anteriormente, pode-se dizer que no geral conseguiu-se obter bons resultados a utilizar esta ferramenta para modelar o processo da digestão anaeróbia.

Para concluir o estudo das RNA realizou-se gráficos tridimensionais para estudar os resultados das previsões, pretendendo assim verificar a capacidade de generalização e o desempenho da melhor rede treinada. Assim sendo, ao analisar estes gráficos que se encontram no anexo VIII, chegou-se à conclusão que a produção de biogás é mais alta quando a temperatura encontra-se nos 38°C, verificando que a rede treinada conseguiu aproximar-se do comportamento do processo da digestão anaeróbia, chegando a ter valores na gama de funcionamento ótimo.

A realização e o desenvolvimento deste estudo ao ser no âmbito de um estágio curricular, representou uma experiência muito gratificante, pois permitiu a integração num ambiente de trabalho, ficando a conhecer a organização e o funcionamento da empresa. Foi uma mais valia pois houve a oportunidade de conhecer de perto o funcionamento da ETAR,

facilitando a compreensão dos processos existentes, especialmente ao processo que interessa para este trabalho, digestão anaeróbia.

Assim este trabalho não fica terminado pois este tema ainda se encontra em desenvolvimento, mostrando que o estudo das RNA é muito importante, pois confirma a sua eficácia na modelação do processo da digestão anaeróbia, permitindo entender o seu funcionamento, conhecer as variáveis que mais impacto têm no processo, com o objetivo de melhorar e otimizar o processo da digestão anaeróbia.

7.2. Sugestões para desenvolvimento futuros

Como referido anteriormente existem variáveis com elevado impacto no processo da modelação das RNA, só pelo facto destas existirem em menor número pode ter condicionado o treino das redes neuronais. Por isso, para trabalhos futuros e se for viável para a empresa recomenda-se aumentar a periodicidade das recolhas para conseguir ter um maior número de amostras destas variáveis. Tendo um conjunto de dados com maior número de casos poderá ajudar para obter melhores resultados no treino das redes.

Outro aspecto que pode influenciar as RNA é a escolha das variáveis, por isso, para futuros estudos poderá ser recomendado a escolha de outras variáveis para comparar e tirar conclusões quais variáveis realmente têm mais impacto para o processo da digestão anaeróbia.

Para além, pode ser feito um tratamento estatístico para complementar as redes neuronais, facilitando a análise dos dados para assim contribuir para melhores resultados, excluindo dados menos fiáveis para o treino das redes.

Posto isto, este trabalho não fica terminado pois este tema encontra-se em desenvolvimento possibilitando a melhoria deste trabalho aqui apresentado. Espera-se que este trabalho contribua para trabalhos futuros dentro do mesmo tema, sendo um ponto de partida para a compreensão do processo da digestão anaeróbia e que contribua para a otimização do processo desta ETAR em estudo bem como outras.

8. Referências Bibliográficas

- Abbasi, T. T. (2012). New York & London.: Springer.
- Águas de Portugal. (18 de Maio de 2016). Obtido de <http://www.adp.pt/pt/>
- Águas do Centro Litoral. (11 de Abril de 2016). Obtido de <http://www.aguasdocentrolitoral.pt/pt/menu-de-topo/atividade/o-sistema/?id=18>
- Alves, M. M. (1998). *Estudo e caracterização de digestores anaeróbios de leito fixo*. Universidade do Minho, Tese de Doutoramento, Braga.
- Appels, L., Baeyens, J., Degreè, J., & Dewil, R. (2008). "Principles and potential of the anaerobic digestion of waste-activated sludge.". *Progress in Energy and Combustion Science* 34 (Elsevier).
- Barreto, J. M. (2002). *Introdução às Redes Neurais Artificiais*. Florianópolis.
- Borja, R., Sanchez, E., & Weiland, P. (1996). Influence of ammonia concentration on thermophilic anaerobic digestion of cattle manure in upflow anaerobic sludge blanket (UASB) reactors. *Process Biochemistry*, 477-483.
- Brandão, J. F. (2014). *OTIMIZAÇÃO DA PRODUÇÃO DE BIOGÁS COM RECURSO A REDES NEURONAIS ARTIFICIAIS*. Instituto Superior de Engenharia do Porto.
- Cardon, A., & Müller, D. (1994). *Introdução às Redes Neurais Artificiais*. Porto Alegre.
- Carmo, J. M. (2013). *Estratégia de Adaptação das ETAR às Alterações Climáticas com Suporte na Simulação Dinâmica*. UNIVERSIDADE DE LISBOA, INSTITUTO SUPERIOR TÉCNICO.
- Carreira, A. C. (2014). *Utilização de Redes Neurais Artificiais na gestão do processo de Digestão Anaeróbia*. Universidade do Minho.
- Cavaleiro, A. J. (1999). *Efeito de Sobrecargas de ácido oleico na actividade de consórcios microbianos anaeróbios desenvolvidos em processos de leito fixo*. Universidade do Minho, Dissertação de Mestrado, Braga.
- CE, C. E. (2001). "Disposal and Recycling Routes for Sewage Sludge.". Luxemburgo: Vols. Parte 3 -Scientific and Technical Report.
- Chen Y., C. J., & S., C. K. (2008). Inhibition of anaerobic digestion process: a review. *Bioresource Technology*, 4044–4064.
- Deublein, D. S. (2008). *Biogas from waste and renewable resources*. Weinheim: Wiley-VCH.
- DL-276/2009. (s.d.).
- Engel, P. M. (s.d.). *Redes Neurais - O ADALINE e o algoritmo LMS*. Obtido de <http://www.inf.ufrgs.br/~engel/data/media/file/cmp121/adaline.pdf>
- Fernandes, M. d. (2015). *Redes Neurais Artificiais na Otimização do Processo de Digestão Anaeróbia - Aplicação a uma ETAR da AdCL*. Instituto Superior de Engenharia do Porto.
- Ferreira, F. (2006). *Modelação e Gestão Integrada de Sistemas de Águas Residuais*. Universidade Técnica de Lisboa, Lisboa.

- Ferreira, F. M. (2012). *SIMULAÇÃO DINÂMICA DO COMPORTAMENTO DE ETAR: PRINCÍPIOS E MODELOS*. Instituto Superior Técnico, Lisboa.
- Forage, R. G. (1985). Effect of environment on microbial activity. *Comprehensive Biotechnology: The Principles, Applications and Regulations of Biotechnology in Industry, Agriculture and Medicine*, M. Moo-Young. (Ed), Pergamon Press Oxford, U.K., 1, 251-280.
- Freire, F. (2003). *Gestão de Energia*. Faculdade de Ciências e Tecnologia da Universidade de Coimbra, Coimbra. Obtido de http://www2.dem.uc.pt/fausto.freire/gestao_energia/_folders/GE_02_03_trab_aluno_pdf/3-biogas.pdf
- Fukuzaki, S. N. (1990). Inhibition of the fermentation of propionate to methane by hydrogen, acetate and propionate. *Environ. Microbiol.*, 719-723.
- Hydromantis. (2006). *GPS-X 5.0 – User's Guide and Technical Reference*. Hydromantis, Inc., Hamilton, Ontario, Canada.
- Leonardo, S. M. (2012). *Caracterização do Digestor Anaeróbio de Lamas Biológicas da ETAR do Choupal, em Coimbra*. Instituto Superior de Engenharia de Coimbra, Coimbra.
- Lettinga, G. (1995). *Anaerobic digestion and wastewater treatment systems*, Antonie van Leeuwenhoek, N.º 67. Kluwer Academic Publishers.
- Lingireddy S. and Brion G. M. (2005). *Artificial Neural Networks in Water Supply Engineering*. ASCE. Reston, Virginia.
- Mara, D., & Horan, N. (2003). *Handbook of water and wastewater microbiology*. London: Academic Press.
- Matos, J. (2008). *Utilização de Redes Neurais Artificiais para a Modelação da Degradação de Sistemas de Distribuição de Água*. Instituto Superior Técnico da Universidade Técnica de Lisboa, Lisboa.
- Mendes, R. (2014). *Contributo para a avaliação da viabilidade do aproveitamento do biogás produzido em digestores anaeróbios a frio para a produção de energia elétrica*. Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa.
- Metcalf&Eddy. (2003). *Wastewater Engineering – Treatment and Reuse*. New York, 4th Edition: McGraw Hill.
- Palisade. (2015). *Add-In de Rede Neural para o Microsoft® Excel. Manual do utilizador. Versão 7*. Palisade Corporation. Ithaca, NY.
- Ribeiro, N. J. (1999). *Estudos Dinâmicos do Processo de Digestão Anaeróbia com vista ao Desenvolvimento de um Sistema Inteligente de Supervisão Baseado em Conhecimento*. Universidade do Minho.
- Rocha, H. T. (2013). *Utilização de Redes Neurais Artificiais na gestão de processos de Digestão Anaeróbia*. Instituto Superior de Engenharia do Porto.
- Silveira, G. (2009). *Estudo de Co-digestão anaeróbia de lamas domésticas com o efluente da indústria da fermentação da levedura do pão da empresa Mauri Fermentos*. Universidade Nova de Lisboa, Dissertação de Mestrado.

- Speece, R. (1983). Anaerobic biotechnology for industrial wastewater treatment. *Environment Science and Technology*, 17(9), pp. 416-427.
- Tchobanoglous, G. B. (1991). *Wastewater engineering*. New York: McGraw-Hill.
- Tchobanoglous, G. B. (2003). *Metcalf & Eddy Wastewater Engineering: Treatment and Reuse*. 4. New York: Mc Graw-Hill.
- Turovskiy, I., & Mathai, P. (2006). *Wastewater sludge processing*. Hoboken. Wiley Interscience.
- Van Lier, J. B. (1995). *Thermophilic anaerobic wastewater treatment; Temperature aspects and process stability*. Tese de doutoramento, Wageningen Agricultural University, Wageningen, Holanda.
- Vavilin, V. A., Rytov, S. V., & Lokshina, L. Y. (1995). Modelling hydrogen partial pressure changes as a result of competition between the butyric and propionic groups of acidogenic bacteria. *Bioresource Technology*, 171-177.
- Vieira, R. C., & Roisenberg, M. (s.d.). *REDES NEURAIS ARTIFICIAIS: UM BREVE TUTORIAL*. Universidade Federal de Santa Catarina (UFSC) – Florianópolis –SC - Brasil .

Anexos

Nota: A acompanhar este trabalho encontra-se um CD onde estão presentes os anexos completos, em ficheiros Excel.

Nestes impressos apenas estão representados alguns exemplos desses ficheiros.

Anexos I

Base de Datos

Tabela 4- Primeiras linhas da base de dados utilizada para treinar as redes neuronais.

Data	T (°C) 1	T (°C) 2	Qentr (m3/dia)	MO (%)	LMMS (%)	LM MO (%)	1 MS (%)	1 pH	AGV 1	Alcali. 1	MO (%)	2 MS (%)	2 pH	AGV 2	Alcali. 2	Qbiogás (m3/dia)
27-fev-15	29,30	29,90	148,60													687,00
28-fev-15	30,05	30,60	104,00													1049,90
02-mar-15	29,90	29,90	161,80													1231,10
03-mar-15	29,40	30,10	194,20													1129,80
04-mar-15	29,35	30,20	104,30													389,00
05-mar-15	29,35	29,95	224,20													431,00
06-mar-15	29,50	30,20	146,60													1508,00
07-mar-15	29,30	31,00	126,00													1418,10
09-mar-15	29,25	30,25	155,80													933,90
10-mar-15	29,70	30,30	133,80													975,40
11-mar-15	29,80	30,30	172,70													1053,50
12-mar-15	29,50	30,00	160,30													763,40
13-mar-15	28,45	30,05	161,10	86,70	5,70	72,80	2,20	7,10	300,00	2600,00	89,80	1,00	7,20	260,00	3000,00	1337,40
14-mar-15	29,00	30,00	212,10													1807,00
15-mar-15	29,25	30,00	184,50													1416,80
16-mar-15	28,90	30,20	206,80													675,20
17-mar-15	28,90	30,20	153,90													2091,50
18-mar-15	30,00	30,00	175,00													1400,90
19-mar-15	28,90	30,05	221,10	79,80	6,65	75,30	2,10	7,20	590,00	2300,00	68,40	2,00	7,20	580,00	2600,00	1522,90
20-mar-15	28,95	29,90	173,60													1527,50
21-mar-15	29,50	29,95	162,90													1633,90
22-mar-15	29,50	30,50	183,50													1689,30
23-mar-15	30,20	29,50	161,00													1521,60
24-mar-15	30,10	30,20	151,20	75,70	6,60	74,20	2,33	7,20	590,00	2500,00	69,90	1,90	7,30	600,00	2800,00	1692,60

Apresenta-se as primeiras linhas da base de dados elaborada para o treino das redes neuronais. A base de dados completa pode ser consultada no ficheiro Anexo I, bem como, as restantes tabelas que foram realizadas para a validação e previsão.

Anexo II

Tabela de Resultados do primeiro conjunto de redes

Tabela 5- Tabela referente aos conjuntos de redes realizados no NeuralTools.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 1	MO (%) 1	23,5127%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	Qentr (m3/dia)	21,4002%		% previsões inadequadas (30% tolerância)	20,0000%	% previsões inadequadas	58,3333%
	MS (%) 1	19,6814%		Erro médio quadrático	231,52	Erro médio quadrático	430,58
	AGV 2	15,2545%		Erro médio absoluto	178,52	Erro médio absoluto	375,09
	MO (%) LM	13,5781%		Desvio padrão de erro absoluto	147,41	Desvio padrão de erro absoluto	211,45
	T (°C) 1	2,0788%					
	pH 1	0,9151%					
	Alcali. 1	0,8685%					
	MO (%) 2	0,7765%					
	Alcali. 2	0,5292%					
	MS(%) LM	0,4588%					
	pH 2	0,4343%					
	MS (%) 2	0,2280%					
	AGV 1	0,2147%					
	T (°C) 2	0,0690%					
RN 2	MS(%) LM	44,2333%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	T (°C) 1	19,5546%		% previsões inadequadas (15% tolerância)	34,0000%	% previsões inadequadas	53,8462%
	pH 1	16,4161%		Erro médio quadrático	239,39	Erro médio quadrático	314,49
	Alcali. 1	13,0245%		Erro médio absoluto	163,15	Erro médio absoluto	243,46
	MO (%) 1	6,7363%		Desvio padrão de erro absoluto	175,18	Desvio padrão de erro absoluto	199,08
	AGV 1	0,0196%					
	MS (%) 1	0,0098%					
	MO (%) LM	0,0034%					
	Qentr (m3/dia)	0,0024%					

Continuação da tabela 5.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN3	AGV 2	16,5492%		% previsões inadequadas (15% tolerância)	24,0000 %	% previsões inadequadas	66,6667 %
	pH 1	12,4671%		Erro médio quadrático	143,80	Erro médio quadrático	338,56
	Qentr (m3/dia)	12,2476%		Erro médio absoluto	96,70	Erro médio absoluto	271,38
	AGV 1	11,5130%		Desvio padrão de erro absoluto	106,43	Desvio padrão de erro absoluto	202,42
	MO (%) 1	11,4473%					
	T (°C) 1	10,0591%					
	pH 2	8,9211%					
RN 4	T (°C) 1	31,5523%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	T (°C) 2	17,8479%		% previsões inadequadas (15% tolerância)	38,0000 %	% previsões inadequadas	53,8462 %
	MO (%) 1	17,6767%		Erro médio quadrático	203,76	Erro médio quadrático	395,38
	MO (%) 2	15,0552%		Erro médio absoluto	139,83	Erro médio absoluto	299,59
	MO (%) LM	12,6039%		Desvio padrão de erro absoluto	148,20	Desvio padrão de erro absoluto	258,01
	Qentr (m3/dia)	3,7655%					
	MS (%) 2	1,4986%					

Continuação da tabela 5.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 5	T (°C) 1	38,7413%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	Alcali. 1	35,1801%		% previsões inadequadas (15% tolerância)	22,0000 %	% previsões inadequadas	84,6154 %
	MO (%) LM	18,5065%		Erro médio quadrático	181,39	Erro médio quadrático	511,04
	MO (%) 1	7,3234%		Erro médio absoluto	126,92	Erro médio absoluto	426,13
	Qentr (m3/dia)	0,2196%		Desvio padrão de erro absoluto	129,59	Desvio padrão de erro absoluto	282,10
	AGV 1	0,0292%					
RN 6	MS(%) LM	51,7122%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	Alcali. 1	37,8733%		% previsões inadequadas (15% tolerância)	42,0000 %	% previsões inadequadas	69,2308 %
	MO (%) 1	9,3790%		Erro médio quadrático	222,27	Erro médio quadrático	404,51
	Qentr (m3/dia)	0,5007%		Erro médio absoluto	165,20	Erro médio absoluto	307,74
	pH 1	0,2122%		Desvio padrão de erro absoluto	148,71	Desvio padrão de erro absoluto	262,53
	T (°C) 1	0,2093%					
	MO (%) LM	0,1132%					

São apresentadas apenas os resultados das primeiras redes treinadas no primeiro conjunto de redes, a tabela completa poderá ser consultada no ficheiro Anexo II.

Anexo III

Exemplo de um relatório gerado pelo NeuralTools

Tabela 6- Tabela gerada do treinamento e teste automático da rede neuronal.

NeuralTools: Treinamento e Teste Automático de rede neural

Executado por: Nádia Amorim

Data: 5 de agosto de 2016 05:18:04

Conjunto de dados: Conjunto de dados núm. 1

Rede: Rede treinada em Conjunto de dados núm. 1 (62)

Resumo	
Informações da rede	
Nome	Rede treinada em Conjunto de dados núm. 1 (62)
Configuração incluída na busca	GRNN, MLFN 2 a 6 nós
Melhor configuração	Previsor numérico de MLFN (2 nós)
Local	Esta pasta de trabalho
Variáveis categóricas independentes	0
Variáveis numéricas independentes	4 (T (°C) 1, Qentr (m3/dia), MO (%) LM, MS(%) LM)
Variável dependente	Var. numérica (Qbiogás (m3/dia))
Treinamento	
Número de casos	50
Tempo de treinamento	02:00:00
Número de tentativas	89160108
Motivo da parada	Parada automática
% previsões inadequadas (15% tolerância)	54,0000%
Erro médio quadrático	252,07
Erro médio absoluto	194,98
Desvio padrão de erro absoluto	159,76
Teste	
Número de casos	13
% previsões inadequadas (15% tolerância)	61,5385%
Erro médio quadrático	358,60
Erro médio absoluto	288,89
Desvio padrão de erro absoluto	212,45
Conjunto de dados	
Nome	Conjunto de dados núm. 1
Número de linhas	398
Tags de casos manuais	NÃO
Análise de impacto de variáveis	
T (°C) 1	34,0218%
Qentr (m3/dia)	33,6438%
MO (%) LM	27,9477%
MS(%) LM	4,3867%

Busca da Melhor Rede			
	Erro de RMS	Tempo de treinamento	Motivo da parada do treinamento
GRNN	418,42	00:00:00	Parada automática
MLFN 2 nós	358,60	02:00:00	Parada automática
MLFN 3 nós	396,37	02:00:00	Parada automática
MLFN 4 nós	420,67	02:00:00	Parada automática
MLFN 5 nós	561,53	02:00:00	Parada automática
MLFN 6 nós	419,19	02:00:00	Parada automática

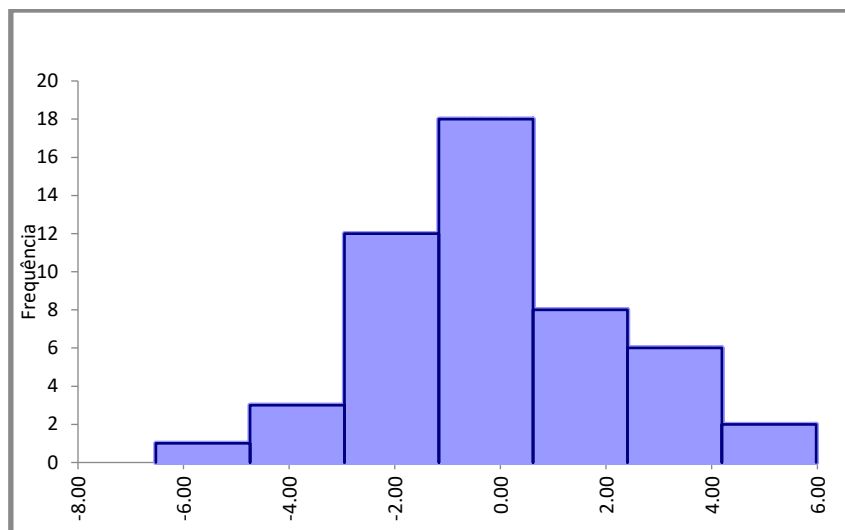


Figura 34- Gráfico da diferença residual, Treinamento.

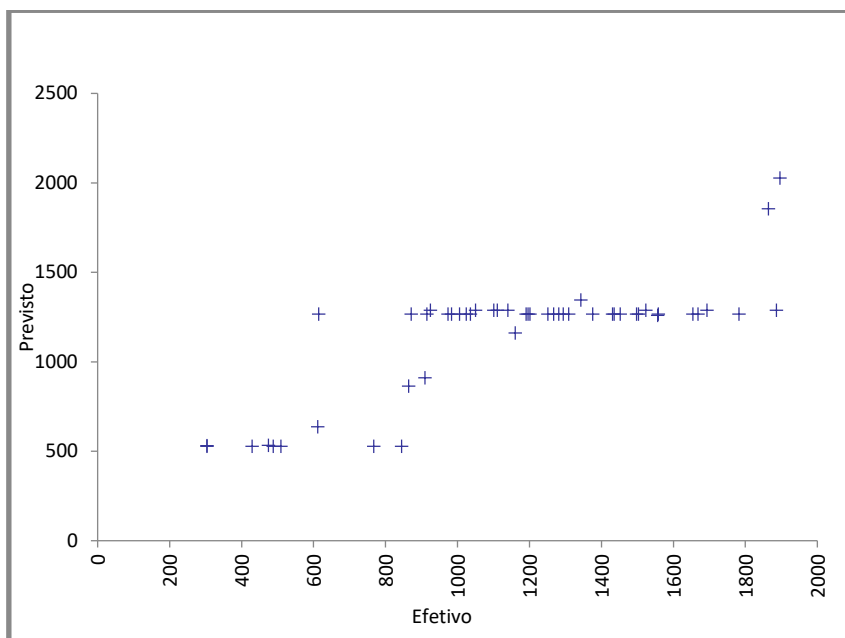


Figura 35- Gráfico dos valores previstos comparados com os valores efetivos, Treinamento.

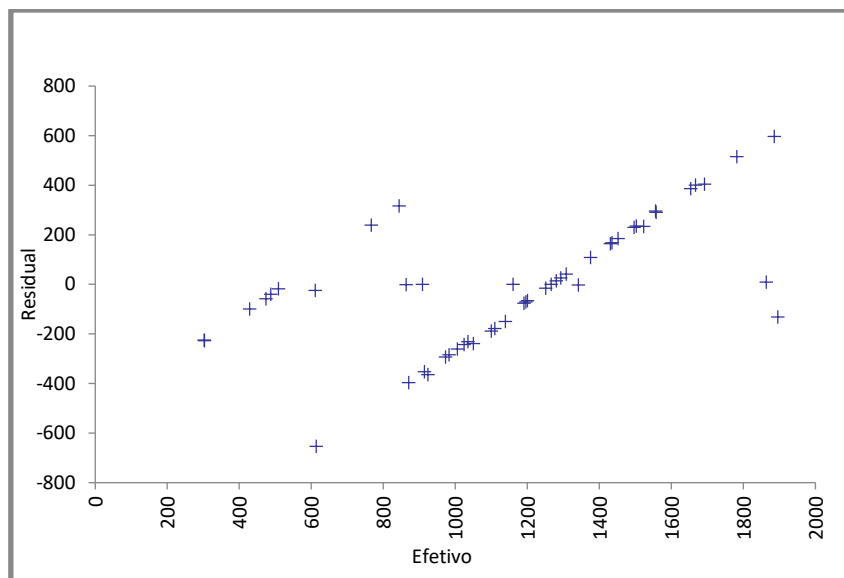


Figura 36- Gráfico com os valores residuais comparados com os valores efetivos, Treinamento.

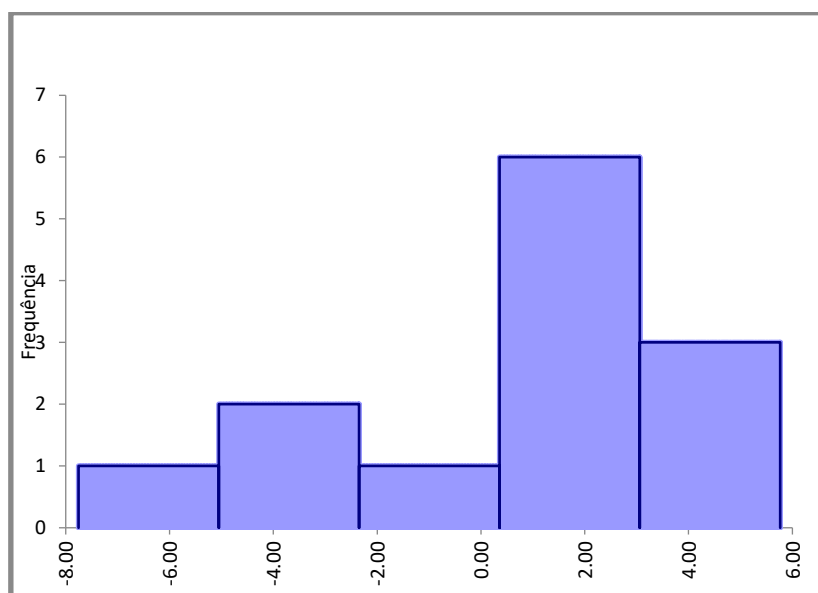


Figura 37- Gráfico das diferenças residuais, Teste.

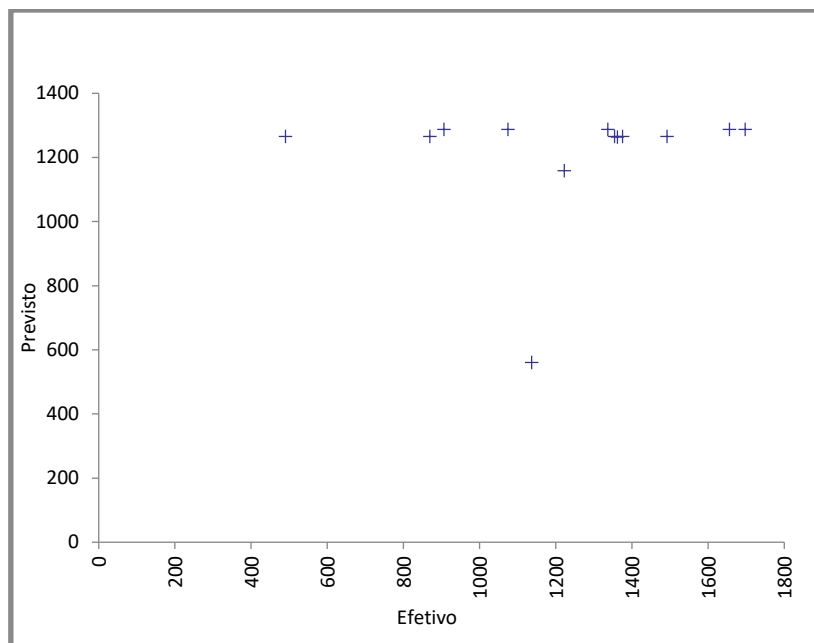


Figura 39- Gráfico dos valores previstos comparados com valores efetivos, Teste.

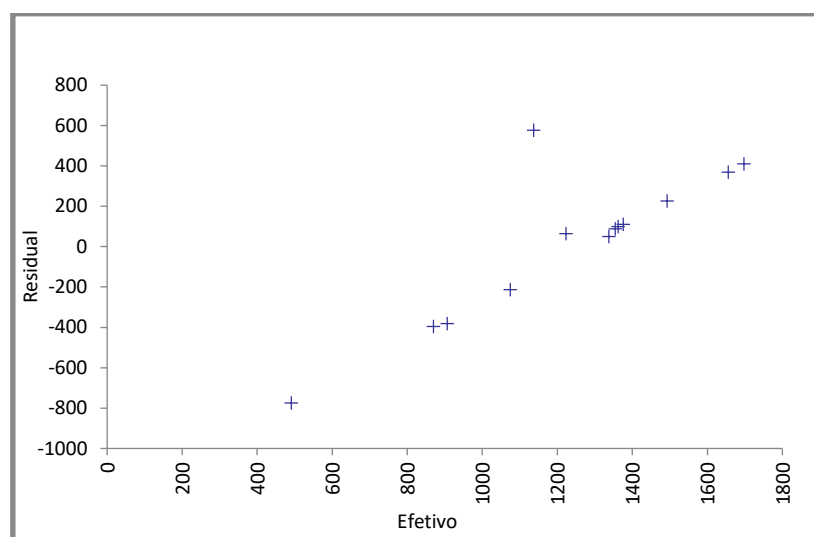


Figura 38- Gráficos com os valores residuais comparados com efetivos, Teste.

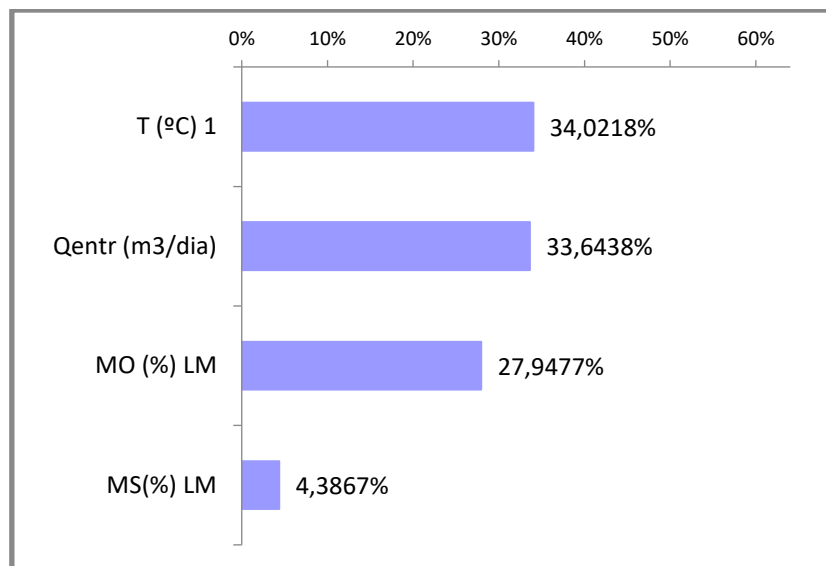


Figura 39- Gráfico do impacto das variáveis.

Anexo IV

Tabela de resultados do segundo conjunto de redes

Tabela 7- Tabela referente aos conjuntos de redes realizados no NeuralTools.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 1	T (°C) 1	44,0499%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	AGV 2	26,1541%		% previsões inadequadas (15% tolerância)	38,0000%	% previsões inadequadas	50,0000%
	MO (%) 1	16,2773%		Erro médio quadrático	230,71	Erro médio quadrático	285,58
	Qentr (m3/dia)	10,9987%		Erro médio absoluto	167,29	Erro médio absoluto	217,56
	MO (%) 2	1,8530%		Desvio padrão de erro absoluto	158,88	Desvio padrão de erro absoluto	185,00
	T (°C) 2	0,6671%					
RN 2	MS(%) LM	41,4628%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	T (°C) 1	29,3202%		% previsões inadequadas (15% tolerância)	24,0000%	% previsões inadequadas	61,5385%
	Alcali. 1	18,4715%		Erro médio quadrático	145,74	Erro médio quadrático	432,53
	pH 1	7,9779%		Erro médio absoluto	99,97	Erro médio absoluto	367,63
	MO (%) 2	2,0493%		Desvio padrão de erro absoluto	106,06	Desvio padrão de erro absoluto	227,88
	Qentr (m3/dia)	0,2838%					
	MS (%) 1	0,2265%					
	MO (%) 1	0,1668%					
	AGV 1	0,0353%					
	MO (%) LM	0,0059%					

Continuação da tabela 7.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 3	MS(%) LM	64,2995%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	Alcali. 1	18,7239%		% previsões inadequadas (15% tolerância)	46,0000 %	% previsões inadequadas	38,4615%
	pH 1	10,7013%		Erro médio quadrático	239,19	Erro médio quadrático	204,91
	T (°C) 1	3,7396%		Erro médio absoluto	189,77	Erro médio absoluto	158,13
	Qentr (m3/dia)	1,1782%		Desvio padrão de erro absoluto	145,60	Desvio padrão de erro absoluto	130,33
	MS (%) 1	0,7725%					
	MO (%) 1	0,5850%					
RN 4	MS(%) LM	73,9051%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	Alcali. 1	16,5443%		% previsões inadequadas (15% tolerância)	48,0000 %	% previsões inadequadas	69,2308%
	MS (%) 1	4,8429%		Erro médio quadrático	261,16	Erro médio quadrático	420,78
	T (°C) 1	2,4751%		Erro médio absoluto	205,84	Erro médio absoluto	328,03
	pH 1	2,1461%		Desvio padrão de erro absoluto	160,73	Desvio padrão de erro absoluto	263,54
	Qentr (m3/dia)	0,0866%					

Continuação da tabela 7.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 5	MS(%) LM	29,2483%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	T (°C) 1	24,0771%		% previsões inadequadas (15% tolerância)	26,0000 %	% previsões inadequadas	69,2308%
	Alcali. 1	23,8086%		Erro médio quadrático	165,18	Erro médio quadrático	400,67
	MS (%) 1	13,2827%		Erro médio absoluto	109,43	Erro médio absoluto	343,37
	MO (%) 1	9,5103%		Desvio padrão de erro absoluto	123,73	Desvio padrão de erro absoluto	206,49
	Qentr (m3/dia)	0,0531%					
	MO (%) LM	0,0199%					
RN 6	MS(%) LM	24,5983%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	Alcali. 1	24,3628%		% previsões inadequadas (15% tolerância)	32,0000 %	% previsões inadequadas	69,2308%
	MS (%) 1	21,6410%		Erro médio quadrático	182,28	Erro médio quadrático	381,58
	T (°C) 1	21,4421%		Erro médio absoluto	120,39	Erro médio absoluto	288,74
	MO (%) 1	7,6591%		Desvio padrão de erro absoluto	136,87	Desvio padrão de erro absoluto	249,45
	Qentr (m3/dia)	0,2968%					

Apresentam-se os primeiros resultados das redes treinadas do segundo conjunto de redes, a tabela completa pode ser consultada no ficheiro Anexo IV.

Anexo V

Tabela dos resultados do terceiro conjunto de redes

Tabela 8- Tabela referente aos conjuntos de dados realizados no NeuralTools.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 1	Qentr (m3/dia)	41,6220%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	MO (%) LM	29,7365%		% previsões inadequadas (15% tolerância)	60,0000 %	% previsões inadequadas	68,7500 %
	pH 1	28,0027%		Erro médio quadrático	312,37	Erro médio quadrático	347,85
	MO (%) 1	0,6388%		Erro médio absoluto	256,15	Erro médio absoluto	293,99
				Desvio padrão de erro absoluto	178,77	Desvio padrão de erro absoluto	185,93
RN 2	Qentr (m3/dia)	74,0675%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	pH 1	20,0122%		% previsões inadequadas (15% tolerância)	33,8462 %	% previsões inadequadas	62,5000 %
	MO (%) LM	5,8744%		Erro médio quadrático	185,86	Erro médio quadrático	353,93
	AGV 1	0,0375%		Erro médio absoluto	141,38	Erro médio absoluto	274,89
	MO (%) 1	0,0085%		Desvio padrão de erro absoluto	120,64	Desvio padrão de erro absoluto	222,94

Continuação da tabela 8.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 3	T (°C) 1	47,0532%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	pH 1	19,1351%		% previsões inadequadas (15% tolerância)	46,0000 %	% previsões inadequadas	53,8462 %
	MO (%) 1	17,4002%		Erro médio quadrático	268,39	Erro médio quadrático	282,06
	Qentr (m3/dia)	16,0799%		Erro médio absoluto	205,44	Erro médio absoluto	216,23
	MO (%) LM	0,3316%		Desvio padrão de erro absoluto	172,70	Desvio padrão de erro absoluto	181,11
RN 4	T (°C) 1	43,5185%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	MO (%) LM	36,3768%		% previsões inadequadas (15% tolerância)	34,0000 %	% previsões inadequadas	53,8462 %
	Alcali. 1	11,9169%		Erro médio quadrático	221,33	Erro médio quadrático	477,68
	Qentr (m3/dia)	7,6902%		Erro médio absoluto	161,76	Erro médio absoluto	363,31
	MO (%) 1	0,4977%		Desvio padrão de erro absoluto	151,07	Desvio padrão de erro absoluto	310,13

Continuação da tabela 8.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 5	Alcali. 1	36,4722%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	MO (%) LM	33,1819%		% previsões inadequadas (15% tolerância)	38,0000 %	% previsões inadequadas	76,9231 %
	T (°C) 1	29,4142%		Erro médio quadrático	215,37	Erro médio quadrático	362,51
	Qentr (m3/dia)	0,9317%		Erro médio absoluto	161,99	Erro médio absoluto	305,59
				Desvio padrão de erro absoluto	141,92	Desvio padrão de erro absoluto	195,01
RN 6	MO (%) 1	19,7024%	Caudal de biogás (m3/dia)	Configuração	MLFN (3 nós)		
	MO (%) LM	18,1189%		% previsões inadequadas (15% tolerância)	20,0000 %	% previsões inadequadas	69,2308 %
	MS (%) 1	17,3822%		Erro médio quadrático	152,24	Erro médio quadrático	385,88
	pH 1	16,9344%		Erro médio absoluto	97,68	Erro médio absoluto	296,27
	T (°C) 1	14,7673%		Desvio padrão de erro absoluto	116,76	Desvio padrão de erro absoluto	247,23
	Qentr (m3/dia)	13,0947%					

Apresenta-se os primeiros resultados das redes treinadas no terceiro conjunto de redes, a tabela completa pode ser consultada no ficheiro Anexo V.

Anexo VI

Tabela de resultados das 3 melhores redes

Tabela 9- Tabela referente aos conjuntos de dados das melhores redes conseguidas pelo NeuralTools.

Nome	Variáveis numéricas independentes	Análise de impacto	Variável numérica dependente	Treinamento		Teste	
RN 11	MO (%) 1	14,4503%	Caudal de Biogás (m3/dia)	Configuração	MLFN (3 nós)		
	T (°C) 1	12,8878%		% previsões inadequadas (15% tolerância)	18,0000%	% previsões inadequadas	61,5385 %
	MS(%) LM	12,6077%		Erro médio quadrático	126,24	Erro médio quadrático	430,36
	MO (%) LM	11,3289%		Erro médio absoluto	77,25	Erro médio absoluto	297,32
	MS (%) 1	10,9526%		Desvio padrão de erro absoluto	99,84	Desvio padrão de erro absoluto	311,14
	pH 1	10,3523%					
	Qentr (m3/dia)	9,8738%					
	Alcali. 1	9,8034%					
	AGV 1	7,7433%					
RN 2	MS(%) LM	41,4628%	Caudal de biogás (m3/dia)	Configuração	GRNN		
	T (°C) 1	29,3202%		% previsões inadequadas (15% tolerância)	24,0000%	% previsões inadequadas	61,5385%
	Alcali. 1	18,4715%		Erro médio quadrático	145,74	Erro médio quadrático	432,53
	pH 1	7,9779%		Erro médio absoluto	99,97	Erro médio absoluto	367,63
	MO (%) 2	2,0493%		Desvio padrão de erro absoluto	106,06	Desvio padrão de erro absoluto	227,88
	Qentr (m3/dia)	0,2838%					
	MS (%) 1	0,2265%					
	MO (%) 1	0,1668%					
	AGV 1	0,0353%					
	MO (%) LM	0,0059%					
RN 9	Qentr (m3/dia)	24,4694%	Caudal de Biogás (m3/dia)	Configuração	MLFN (3 nós)		
	MS(%) LM	21,4593%		% previsões inadequadas (15% tolerância)	26,0000%	% previsões inadequadas	61,5385 %
	MS (%) 1	21,4252%		Erro médio quadrático	146,20	Erro médio quadrático	542,19
	T (°C) 1	17,2326%		Erro médio absoluto	112,01	Erro médio absoluto	415,74
	MO (%) LM	15,4134%		Desvio padrão de erro absoluto	93,95	Desvio padrão de erro absoluto	348,04

Apresenta-se apenas os resultados das 3 melhores redes encontradas dos conjuntos realizados, as tabelas completas das restantes redes também com bons resultados poderão ser consultadas no ficheiro Anexo VI.

Anexo VII

Dados e Gráfico de interpretação dos dados da validação

Tabela 10- Base de dados referentes aos valores utilizados para a validação.

Relatório de Previsão: "Rede treinada em Conjunto de dados núm. 1 (62)"

Data	T (°C) 1	Qentr (m3/dia)	MO (%) LM	MS(%) LM	Qbiogás (m3/dia)	Tag usada	Previsão
06-mai-16	31,00	68,10	80,10	7,90	606,70	prever	527,17
12-mai-16	32,40	76,60	76,70	9,60	609,80	prever	527,17
17-mai-16	33,00	78,70	76,00	8,10	743,1	prever	527,17
19-mai-16	33,15	55,10	74,40	8,90	907,20	prever	527,17
03-jun-16	34,25	70,50	75,90	7,90	914,70	prever	527,17
08-jun-16	34,85	179,20	77,30	8,20	877,80	prever	527,90
14-jun-16	32,25	154,50	76,00	8,10	743,00	prever	1287,93
17-jun-16	31,30	152,90	69,00	7,90	301,40	prever	1287,93
23-jun-16	31,00	182,90	79,30	8,20	1429,60	prever	1287,93

A partir da tabela de cima elaborou-se este gráfico para uma melhor interpretação de dados.

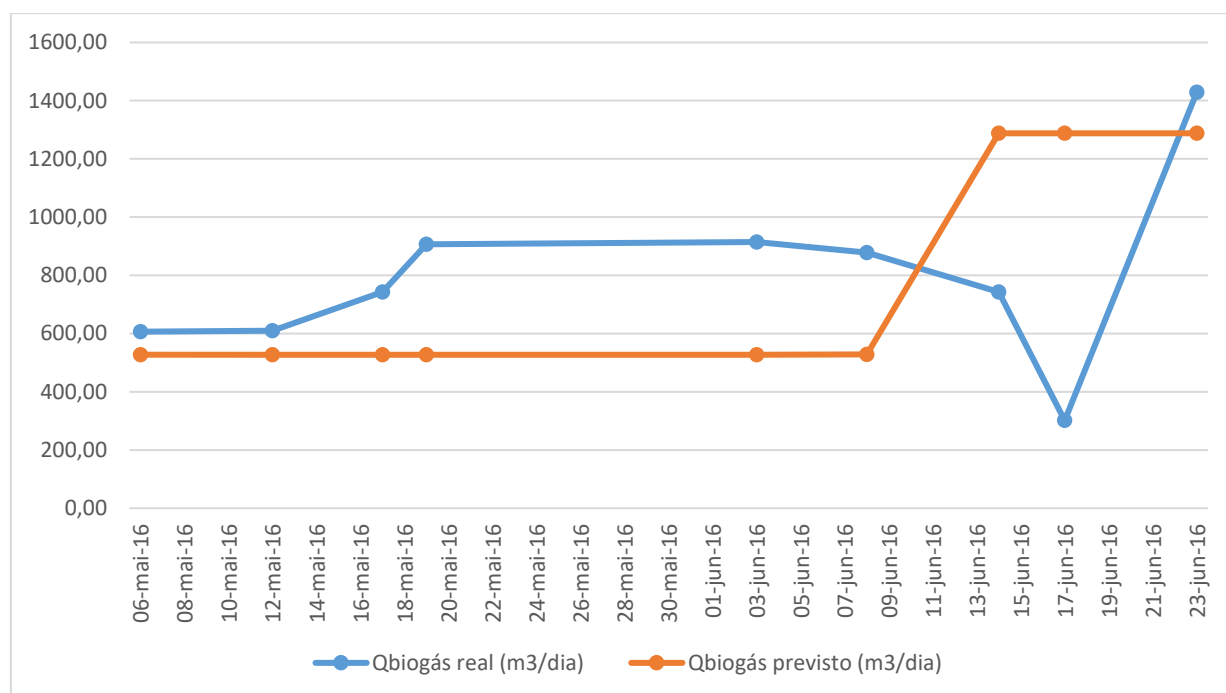


Figura 40- Gráfico de validação da melhor rede.

Anexo VIII

Dados e gráficos de interpretação dos dados de previsão

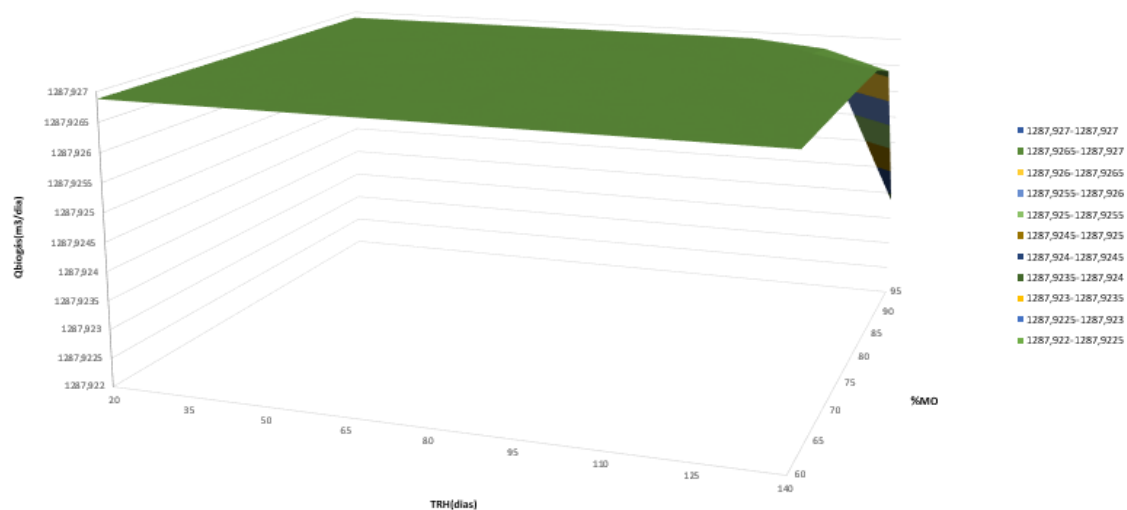


Figura 41- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 28°C.

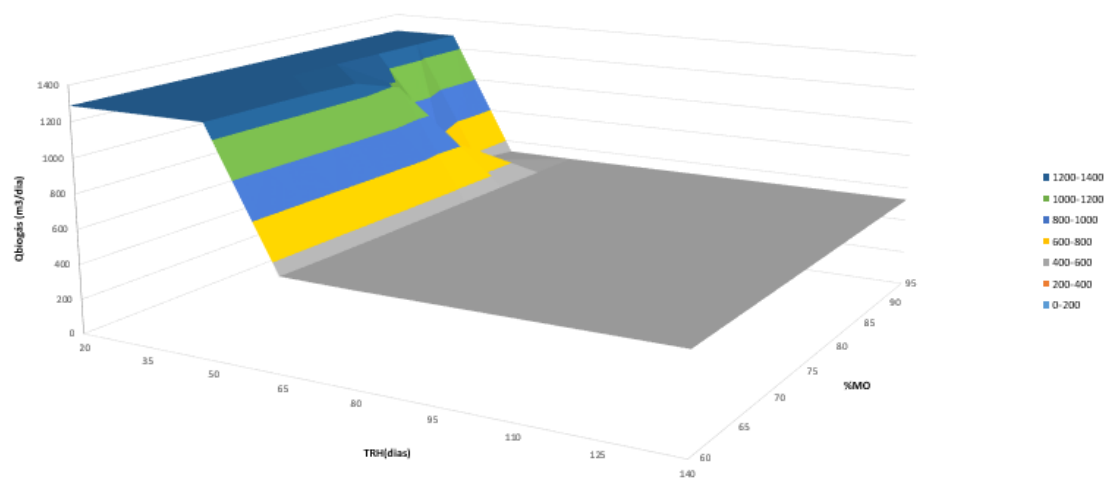


Figura 42- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 32°C.

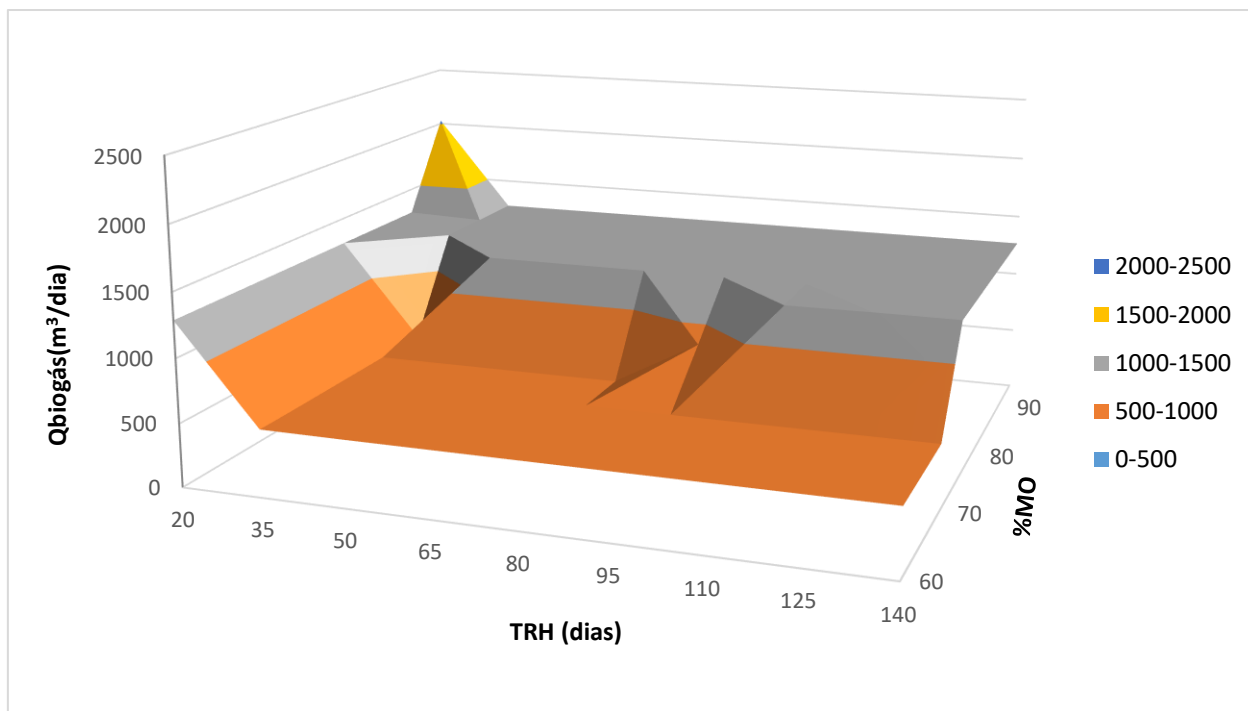


Figura 43- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 35°C.

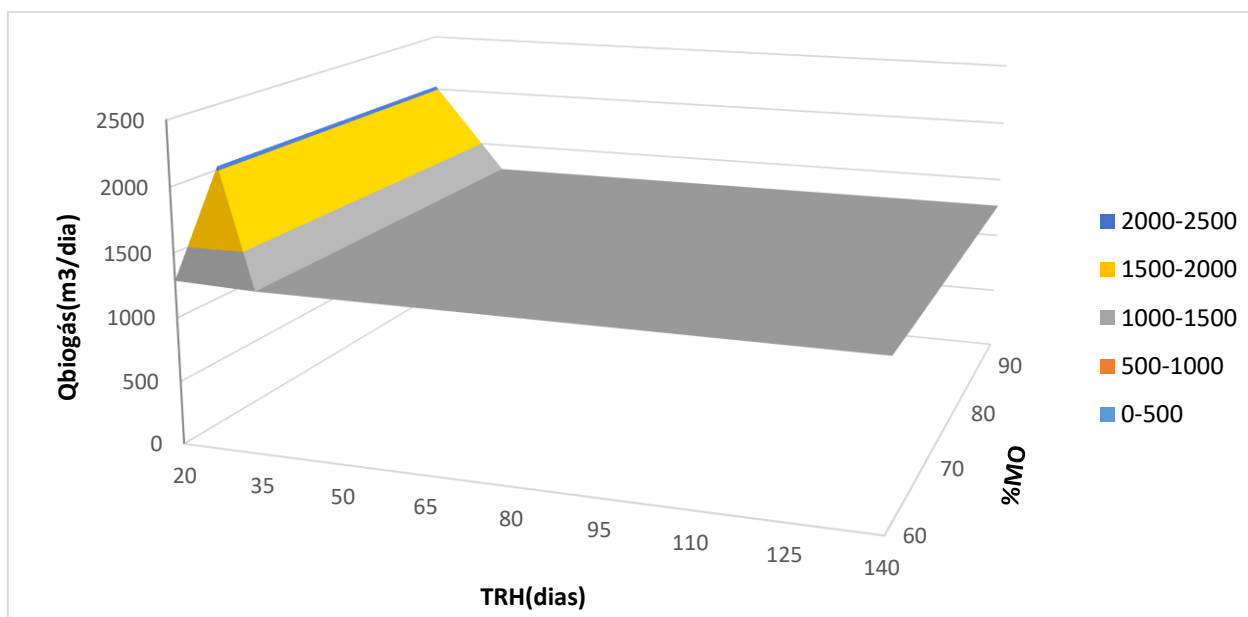


Figura 44- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 38°C.

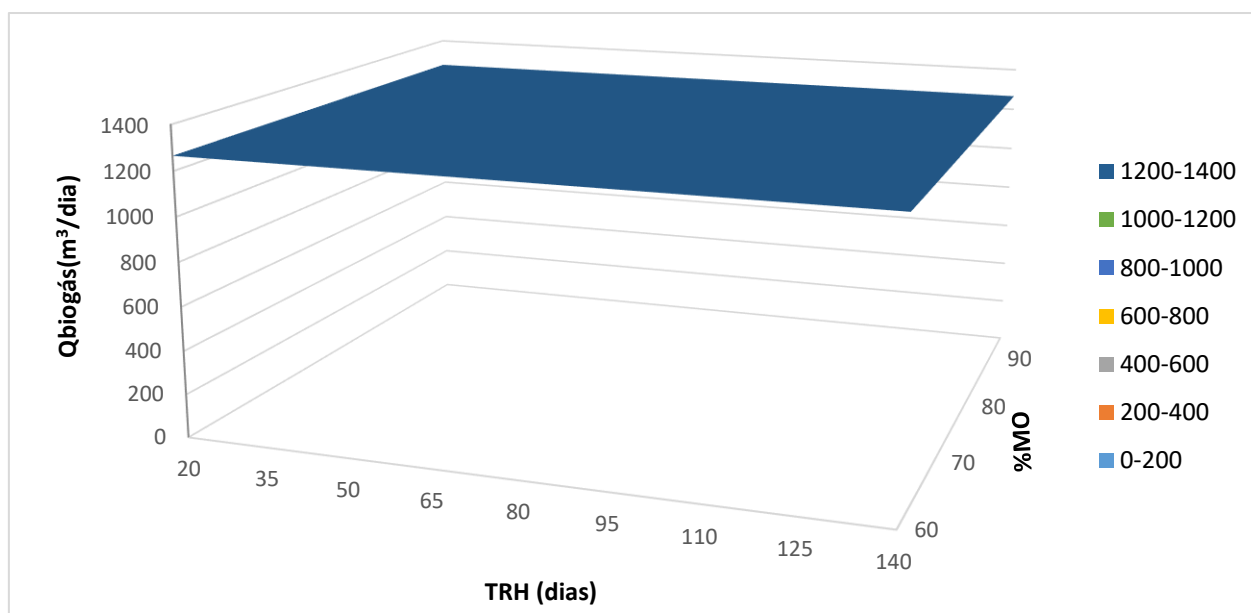


Figura 45- Gráfico referente à interpretação dos dados de previsão dos valores relativos à temperatura de 42°C.